Ville Mäntyniemi

Sound Synthesis in Working Machine Simulators

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 21.10.2014

Thesis supervisor:

Prof. Vesa Välimäki

Thesis advisor:

D.Sc. (Tech.) Rémi Mignot

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Ville Mäntyniemi

Title: Sound Synthesis in Working Machine Simulators

Date: 21.10.2014 Language: English Number of pages: 8+60

Department of Signal Processing and Acoustics

Professorship: Audio Signal Processing Code: S-89

Supervisor: Prof. Vesa Välimäki

Advisor: D.Sc. (Tech.) Rémi Mignot

This thesis aims to improve the sound environment in working machine simulators. Modern simulators are visually and operationally extremely advanced and realistic, but the sound environment is still limited. By improving the sounds in these simulators, the simulators can grow into exceptionally realistic training tools with the ability to fully educate future operators in a completely safe environment.

Existing sounds are improved and new sounds are created for three different simulator types: a forest harvester and forwarder simulator, a drill rig simulator, and a truck-mounted hydraulic platform simulator. The main sound types synthesized are hydraulic sounds, drilling sounds, feeding and delimbing sounds (forest machines), and basic contact sounds.

Linear predictive coding is used throughout the project in synthesizing several different sounds, including hydraulic sounds and harvester sounds. Spectral subtraction is also widely utilized in creating noiseless contact sound samples required in the forwarder simulator. Other methods used include filtering, amplitude envelope extraction, and peak detection.

Keywords: acoustics, drilling, forestry machines, hydraulics, simulators

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Ville Mäntyniemi

Työn nimi: Työkonesimulaattoreiden äänisynteesi

Päivämäärä: 21.10.2014 Kieli: Englanti Sivumäärä: 8+60

Department of Signal Processing and Acoustics

Professuuri: Audiosignaalinkäsittely Koodi: S-89

Valvoja: Prof. Vesa Välimäki

Ohjaaja: TkT Rémi Mignot

Tässä työssä pyritään parantamaan työkonesimulaattoreiden äänisynteesiä. Nykypäivän simulaattorit ovat visuaalisesti ja toiminnaltaan erittäin kehittyneitä, mutta niiden äänimaailma on vielä realistisuuden kannalta puutteellinen. Parantamalla simulaattoreiden ääniympäristöä simulaattorit voivat ottaa suuria kehitysaskelia ja saavuttaa roolin täysin realistisina opetustyökaluina, joiden avulla tulevat koneenkuljettajat voidaan kouluttaa täysin realistisessa, mutta turvallisessa ympäristössä.

Työssä parannetaan simulaattoreiden olemassaolevia ääniä sekä luodaan täysin uusia ääniä kolmelle eri simulaattorityypille: metsäkonesimulaattori, porauskonesimulaattori sekä nostolavalaitesimulaattori. Oleellisimmat syntetisoidut äänet tässä projektissa ovat hydrauliset äänet, porausäänet, syöttö- ja karsintaäänet (metsäkoneet) sekä kontaktiäänet.

Työssä hyödynnetään useita eri signaalinkäsittely- sekä äänisynteesimenetelmiä, kuten lineaariprediktiota, suodatusta, spektrivähentämistä ja huipuntunnistusta. Lineaariprediktiota käytetään äänisynteesissä huomattavan paljon työn aikana, erityisesti hydraulisten äänien ja metsäkoneäänien parissa.

Avainsanat: akustiikka, hydrauliikka, metsäkoneet, poraaminen, simulaattorit

Preface

First, I would like to thank the Finnish Work Environment Fund (TSR), Creanex, Sandvik, and Aalto for funding this project. My supervisor, Prof. Vesa Välimäki, I would like to thank for the opportunity to be a part of this interesting venture. It has been an extremely invaluable experience for me and I have learned a great deal. Thank you also to my instructor, D.Sc. Rémi Mignot, for the feedback and time spent together on this project. Thank you also to all the people at Creanex and Sandvik, who made this project an enjoyable one. Lastly, I would of course like to thank my family and friends for all their love and support.

Otaniemi, 21.10.2014

Ville Mäntyniemi

Contents

Symbols and abbreviations

Symbols

a_k	predictor coefficients
A	transfer function of the inverse filter
β	ratio of the main lobe width and side lobe levels in a Kaiser window
e, E	residual
E_{MSE}	mean square error
E_p	minimum squared error
F_1	state variable filter tuning coefficient
f_c	cutoff frequency
f_s	sampling frequency
G	gain
H	transfer function of the spectral shaping filter
j	imaginary unit
n	sample
n_k, N_k	noise signal
N	filter order
p	LP filter order
Q_1	state variable filter tuning coefficient
R	autocorrelation function
s, S	signal
S_A	magnitude averaging output
t	time
θ	phase
u_n, U	excitation signal
μ	average
w	Kaiser window function
ω	angular frequency
x, X	input signal
y	output signal
y_b	band-pass output signal
y_h	high-pass output signal
y_l	low-pass output signal
z	Z-transform variable

Operators

$\frac{\delta y}{\delta x}$	derivative of y in relation to x
$\sum\limits_{k=m}^{p}$	sum with lower bound m and upper bound p
$\hat{x}$	estimate of x
$\overline{x}$	average of x

Abbreviations

AR	autoregressive
ARMA	autoregressive moving average
FFT	fast Fourier transform
FIR	finite impulse response
LP	linear prediction
LPC	linear predictive coding
MA	moving average
MSE	mean square error
STFT	short time Fourier transform

1 Introduction

Training simulators are rapidly becoming essential tools in the training of new machine operators. These simulators allow for a completely safe training environment, where future operators can efficiently learn the essential operations, tasks, and hazards. Ideally, a simulator-trained operator will be able to correctly and safely operate an actual machine without prior real-life experience. To reach this goal, the simulator environment must be as realistic as possible to minimize the transition between the simulator and the machine, cf. e.g. [1].

Although modern simulators are already visually and operationally extremely realistic and advanced (Fig. 1), the sound environment is still in need of considerable improvement. Sounds play a crucial role in operating several machines: they give the operator valuable information about the state of the machine, they inform of other machines in the area, and they give audible cues which help in operating the machines. Essentially, with a realistic sound environment the operators can be trained to distinguish defects in the machine, avoid hazardous situations, and operate the machines with the help of both visual and aural aids. This method of training will allow beginners to learn these aforementioned skills in a hazard-free environment instead of an actual worksite with real machines and people. Although simulator based sound synthesis work does exist relating to flight simulators [2][3], little to no research can be found regarding working machine simulators, excluding drilling sounds [4].

Figure 1: Working machine simulators: forest machine simulator (left) and drill rig simulator (right). Pictures retrieved from www.creanex.com.

This thesis will focus on improving the sound environment of three different types of working machine simulators: a forestry harvester and forwarder simulator, a drill rig simulator, and a truck mounted hydraulic platform simulator. The actual machines in question are presented in Fig. 2. The method of improving the sounds in these simulators will be to improve the existing sounds and to add completely new sounds which are currently missing, but necessary for a more realistic soundscape. The current sound synthesis method used in the simulators is sampling synthesis [6], with some basic processing such as pitch shifting and amplitude modulation. Sampling synthesis will remain as the basis for the whole sound system, i.e. sound files will be played back in the simulator when they are required, but other methods of sound synthesis will be implemented to improve the authenticity of these sounds and to add new sounds. Subtractive synthesis [6] will play a major role in synthesizing several different sounds in this project, as one common method falling under this category is Linear Predictive Coding (LPC) [7][8], which is a vital tool widely utilized in this project.

Figure 2: Top left: forest harvester, top right: forest forwarder, bottom left: truck mounted hydraulic platform, bottom right: drill rig. Pictures retrieved from www.ponsse.com, www.miningandconstruction.sandvik.com, and www.bronto.fi.

The structure of this thesis is as follows. Section 2 begins with an introduction to sound synthesis and the theory behind the methods used to synthesize the sounds created in this project. Section 3 explains the synthesis process behind creating hydraulic sounds, which were missing from the simulators completely. Hydraulic sounds can be very prominent in the actual machines and they exist in both the forestry machines and drill rigs, making them a vital part of this thesis. Section 4 focuses on drilling sound synthesis, the most essential sound in a drill rig. Although synthesis work based on physical models can be found related to drilling sounds [4][5], this thesis will focus on improving the existing sample based drilling sound and not utilize the more advanced physical models. Section 5 deals with sound synthesis related to the forest machine simulators. Important sounds synthesized in this section include, feeding, delimbing, and contact sounds. Section 6 presents the results of a listening test regarding the hydraulic and forest simulator sounds. The listening test was carried out to obtain valuable feedback and comments on the synthesized sounds to help further improve the sound environment. Finally, Section 7 summarizes the project and discusses future work.

2 Background & Methods

2.1 Sound Synthesis

Today sound synthesis is generally divided into four different classes: processed recordings, spectral models, physical models, and abstract algorithms [9]. This project will focus mostly on synthesis in the processed recordings and spectral model classes. Synthesis methods falling under these two categories will be explained in more detail below.

2.1.1 Sampling and Wavetable Synthesis

As explained in [6], sampling and wavetable synthesis are currently the most popular methods used in sound synthesis and they are both examples of processed recordings. In sampling synthesis, recorded sounds are simply played back with or without processing. An example of this is a digital piano, where recordings of different piano keys being played are used to synthesize a piano sound. There are several recordings for each key to simulate different attacks, which are needed to create a realistic sounding digital piano. Sampling synthesis is the main method used in the simulators to synthesize sound prior to the start of this project. Recorded machine sounds were played backed in the simulators with some simple processing, including amplitude variations and pitch shifting.

In wavetable synthesis a single period of a sound waveform is stored into a table and then repeated. For musical instruments it is common to store separate parts of the sound. For example, the attack, sustain, and release segments of the wanted sound can be stored separately and then played back when needed to create a more realistic synthesis result. Filters are typically used with wavetable synthesis to achieve variation in sound through spectral control. In addition to this, samples may be pitch shifted to a certain degree to allow for more flexibility. If a sample is pitch shifted too much, it will start to sound unnatural. Instead several different samples with different frequencies should be used and then pitch shifting should implemented between these samples to obtain the required frequency. This method is called multisampling and it can be used with sounds created in this project.

2.1.2 Additive and Subtractive Synthesis

Two examples of spectral modeling are additive and subtractive synthesis. Additive synthesis is simply the method of combining several sine waves to create a sound. In subtractive synthesis a spectrally rich waveform is filtered with a suitable filter to achieve the required sound. The excitation signal is commonly white noise or an impulse train and it is filtered with a filter which will shape the spectrum in a required way. One common subtractive synthesis method is Linear Predictive Coding (LPC), which can be used to extract the spectral shape of a signal and design a suitable filter to model this spectrum. LPC is used extensively in this project, especially in synthesizing hydraulic sounds. Linear predictive coding will be explained in more detail in the next section.

2.2 Linear Predictive Coding

2.2.1 Basic Principle

Linear predictive coding is a common technique used in audio signal processing and especially speech processing, cf. [7][8]. It is used to present the spectral characteristics of a signal. LPC assumes signals to present a source-filter model, where a source is excited by a linear filter, e.g. in speech processing the source is the vocal cords and the filter models the vocal tract. The LPC filter models the spectral shape of the sound produced by the vocal tract and a white noise excitation can be used to synthesize a speech signal. This is called LPC synthesis and it is widely utilized in this project, although not for its most common application, speech processing, but instead for machine sound synthesis. LPC synthesis will be explained later in more detail.

A basic signal can be written as

$$S(z) = U(z)H(z), \tag{1}$$

where z is the Z-transform variable and $U(z)$ is the Z-transform of the excitation signal u_n, which is filtered by the transfer function of the spectral shaping filter $H(z)$, which is an estimate of the spectral shape of the signal $s(n)$. For continuous spectra, the excitation $U(z)$ is assumed to have a flat magnitude spectrum. In this project white noise is almost exclusively used as the excitation signal.

There are three different cases of the LPC-model: the all-pole model (autoregressive = AR model), the all-zero model (moving average = MA model), and the pole-zero model (autoregressive moving average = ARMA model), cf. [8]. The all-pole model is the most widely used and will be used exclusively in this project. In the all-pole model, the signal $s(n)$ is a linear combination of past values with an input u_n and gain G

$$s(n) = \sum_{k=1}^{p} a_k s(n-k) + G u_n. \tag{2}$$

The transfer function of the all-pole model is given as

$$H(z) = \frac{G}{A(z)} = \frac{G}{1 - \sum\limits_{k=1}^{p} a_k z^{-k}}. \tag{3}$$

Linear prediction must be used to determine the predictor coefficients a_k of the FIR filter $A(z)$, which is commonly known as the inverse filter.

2.2.2 Linear Prediction

Linear prediction is the method of predicting future samples by forming estimates from linear combinations of previous samples. The linear predictor is shown in Eq.(4), where $\hat{s}(n)$ is the estimate, $s(n-k)$ the previous sample, and a_k the predictor coefficients. The amount of previous samples used for the linear prediction is presented by p, which is the order of the inverse filter. The idea behind linear prediction is to calculate the predictor coefficients a_k, so that the difference between the estimate $\hat{s}(n)$ and the actual sample $s(n)$ is as small as possible [6].

$$\hat{s}(n) = \sum_{k=1}^{p} a_k s(n-k) \tag{4}$$

The error of the estimate is the difference between the actual sample and the estimate, which is also known as the residual

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^{p} a_k s(n-k). \tag{5}$$

In linear prediction, the predictor coefficients are obtained by minimizing the Mean Square Error (MSE)

$$E_{MSE} = \sum_{n} e(n)^2 = \sum_{n} \left(s(n) - \sum_{k=1}^{p} a_k s(n-k) \right)^2 \tag{6}$$

by taking the derivative in relation to the predictor coefficients and setting it to zero

$$\frac{\delta E}{\delta a_i} = 0, \qquad 1 \leq i \leq p \tag{7}$$

Then we arrive at

$$\sum_{k=1}^{p} a_k \sum_{n} s(n-k)s(n-i) = \sum_{n} s(n)s(n-i), \qquad 1 \leq i \leq p. \tag{8}$$

By expanding Eq.(6) and substituting Eq.(8), the minimum total squared error, E_p, can be obtained

$$E_p = \sum_{n} s_n^2 + \sum_{k=1}^{p} a_k \sum_{n} s_n s_{n-k}. \tag{9}$$

After this step, there are two common choices for calculating the predictor coefficients: the autocorrelation method and the covariance method [8]. The autocorrelation method is used in this project and is explained below.

2.2.3 Autocorrelation Method

In the autocorrelation method, the error in Eq.(6) is minimized within the boundaries $-\infty < n < \infty$ (in practice, windowing is applied), which reduces Eqs.(8) and (9) to

$$R(i) = \sum_{k=1}^{p} a_k R(i-k), \qquad 1 \leq i \leq p \tag{10}$$

and

$$E_p = R(0) - \sum_{k=1}^{p} a_k R(k). \tag{11}$$

The optimal LPC shown in Eq.(10) is presented below in matrix form

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ R(1) & R(0) & \ddots & \vdots \\ \vdots & & \ddots & R(1) \\ R(p-1) & \dots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}. \tag{12}$$

Because the left side of the matrix equation is a Toeplitz matrix (each descending diagonal from left to right is constant), the predictor coefficients can be recursively obtained using the Levinson-Durbin Algorithm [8]

$$E_0 = R(0) \tag{13}$$

$$k_i = \left[R(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) \right] \Big/ E_{i-1} \tag{14}$$

$$a_i^{(i)} = k_i \tag{15}$$

$$a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}, \qquad 1 \leq j \leq i-1 \tag{16}$$

$$E_i = (1 - k_i^2) E_{i-1}. \tag{17}$$

Equations (14) - (16) are recursively calculated with $i = 1, 2, 3, ...p$. From this algorithm we can then obtain the optimal predictor coefficients for the inverse filter

$$A(z) = 1 - \sum_{k=1}^{p} a_k z^{-k}. \tag{18}$$

As presented in [8], the gain can be written as

$$G = \sqrt{E_p} = \sqrt{R(0) - \sum_{k=1}^{p} a_k R(k)} \tag{19}$$

2.2.4 LPC Synthesis

Linear prediction can be considered as the process of dividing a signal into two parts: the inverse filter $A(z)$ and the residual $e(n)$. The opposite process of this analysis method is LPC synthesis, where the original signal can be obtained by filtering the residual with the inverse of $A(z)$, also known as the LPC filter

$$S(z) = E(z) \frac{1}{A(z)} = \frac{E(z)}{1 - \sum_{k=1}^{p} a_k z^{-k}}. \tag{20}$$

The LPC filter models the spectral shape of the signal and it can be effectively applied in sound synthesis. Instead of filtering the residual of the original signal, a suitable excitation signal is selected (usually white noise), which is then filtered with the LPC filter. The output is a synthetic sound signal with the spectral qualities of the original signal. Due to the noisy properties of the residual, replacing it with a synthesized noise signal is a perceptually valid method. By using sufficiently large filter orders p the difference between the signals may be inaudible.In addition, increasing the filter order p greatly augments the detail of the spectral envelope, which is why high filter orders ($p = 1000$) are preferred throughout this project. The LPC synthesis method is applied throughout this project.

2.3 Filtering

Filtering is an essential tool in signal processing and audio signal processing, so it is no surprise it is also widely practiced in this project. Basic FIR-filters (finite impulse response filters) are applied throughout the project, including low-pass, band-pass, band-stop, and high-pass filters [12]. In addition to these basic filters, Moog filters [14] [15], state variable filters, and sliding average filters [13] are also utilized in the processing of certain sounds.

2.3.1 FIR-filter Design

The FIR-filters are designed using the classic window design method [17] with a Kaiser window [10]. In the window design method, the specifications of the filter are determined in the frequency domain by the wanted cutoff frequencies of the passband and stopband and their corresponding maximum ripple values. The filter is then designed to meet these specifications using a specific window function. If no window function is used, i.e. the window is a rectangular window, there will always be passband ripple, which can be observed in Fig. 3.

As seen in Fig. 3, passband ripple can be easily minimized by using a nonrectangular window, e.g. Hamming, Hanning, Chebyshev, or Kaiser window. In this case, a Kaiser window will be implemented due to its flexibility in allowing the user to control the width and levels of the main and side lobes. The Kaiser window is given as

$$w(k) = \frac{I_0\left[\beta\sqrt{1-\left(\frac{k-p}{p}\right)^2}\right]}{I_0(\beta)}, \qquad k = 0,1,2,...,N-1 \text{ and } p = (N-1)/2, \quad (21)$$

where, I_0 is the zeroth order modified Bessel function of the first kind, cf. [11]

$$I_0(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{(k!)^2} \quad (22)$$

The parameter β controls the ratio of the main lobe width and side lobe levels: as β is increased, the main lobe width increases and the side lobe amplitudes decrease.

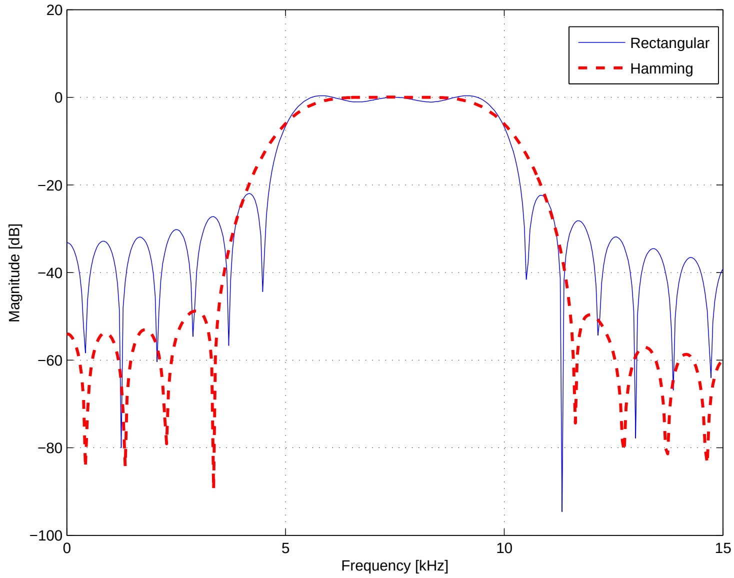

Figure 3: Example spectrum of a band-pass filter with a passband of $[5, 10]$ kHz using a rectangular window and a Hamming window.

The optimal solution is to try to reduce the sidelobe amplitudes without increasing the main lobe width too much [17]. Three different Kaiser windows are shown in Fig. 4. Note how there is no passband ripple and how increasing β affects the main and side lobes.

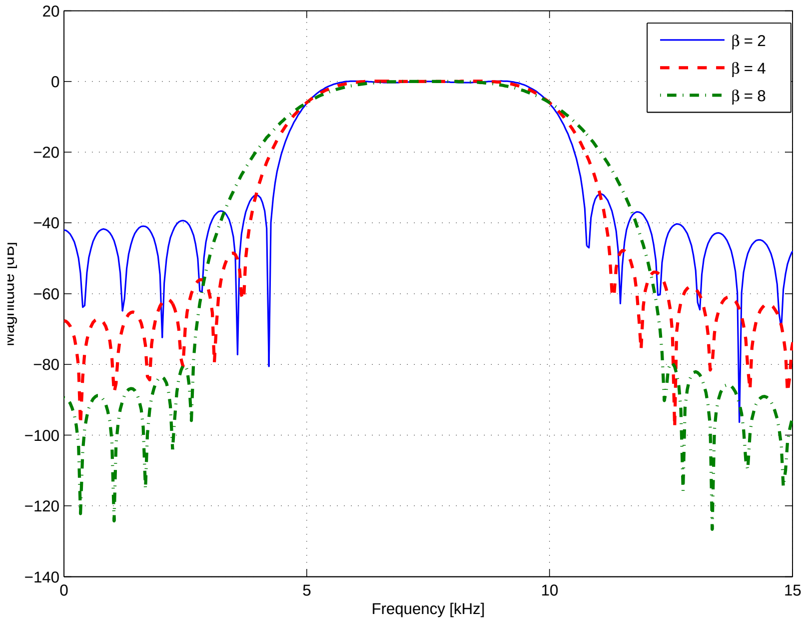

Figure 4: Three different Kaiser windows with different values of β with a passband of [5,10] kHz.

2.3.2 Sliding Average Filter

The sliding average filter, also known as a moving average filter, is a very useful FIR filter. The output of the filter is the average of a number of points determined by the window size N

$$y(k) = \frac{1}{N} \sum_{i=0}^{N-1} x(k-i). \tag{23}$$

A whole set of data can be processed by sliding the window through the data, so that an average will be calculated for each sample from N neighboring samples. This process efficiently smooths out the data removing any short-time fluctuations and emphasizing long-term changes. In signal processing, this process can be considered low-pass filtering and it can be used to filter out noise or to extract the amplitude envelope of a signal, cf. [13]. Amplitude envelope extraction is what the sliding average filter will be used for in this project.

The amplitude envelope of a signal is extracted by first performing a full wave rectification, after which it is filtered with the sliding average filter. The window size N determines how much smoothing is applied to the signal, i.e. how accurately all the valleys and peaks are featured. Increasing the window size increases the smoothness of the amplitude envelope and less features of the original signal will be present. Figure 5 presents the process explained above.

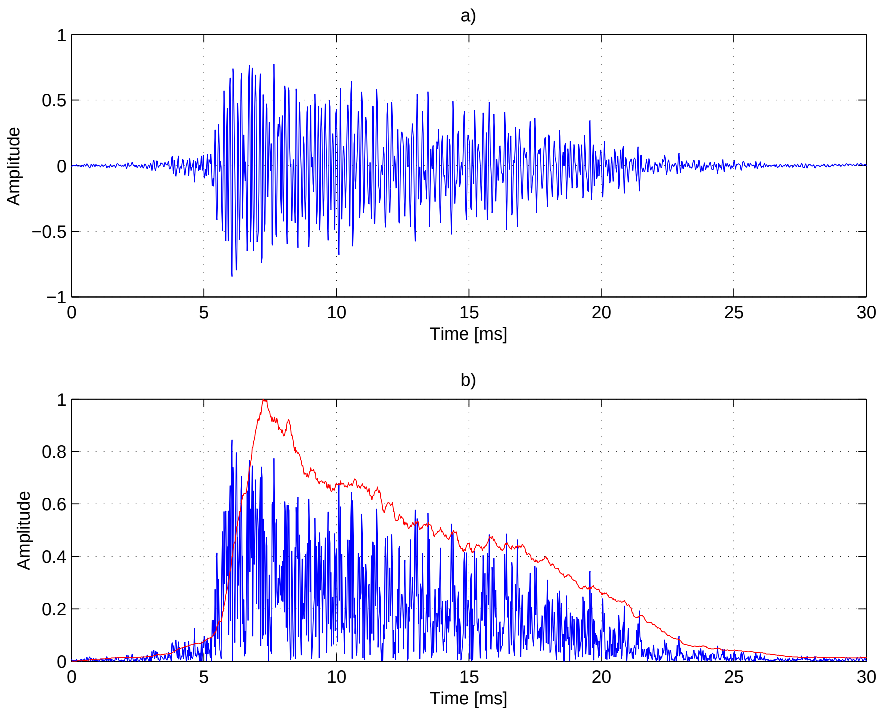

Figure 5: a) Original signal and b) full-wave rectified signal with amplitude envelope in red.

2.3.3 Moog Filter

The Moog filter is a time varying filter commonly used in musical applications such as synthesizers, effects, and samplers. Originally published as a voltage-controlled filter by Robert Moog in 1965 [14], the Moog filter has since been converted into a non-linear digital implementation by Huovilainen [15], which included five nonlinear functions inside the filter sections. A more efficient single nonlinearity version of the digital Moog filter implementation is presented by Välimäki and Huovilainen in [16]. This version of the Moog filter implementation will be applied in the processing of certain sounds during the project. Essentially, the filter is able to slide the cutoff frequency of a low-pass filter with a resonance from one frequency to another (Fig. 6). This feature is excellent for creating a sweeping-type sound effect, where the frequency characteristics of a sound change in relation to time.

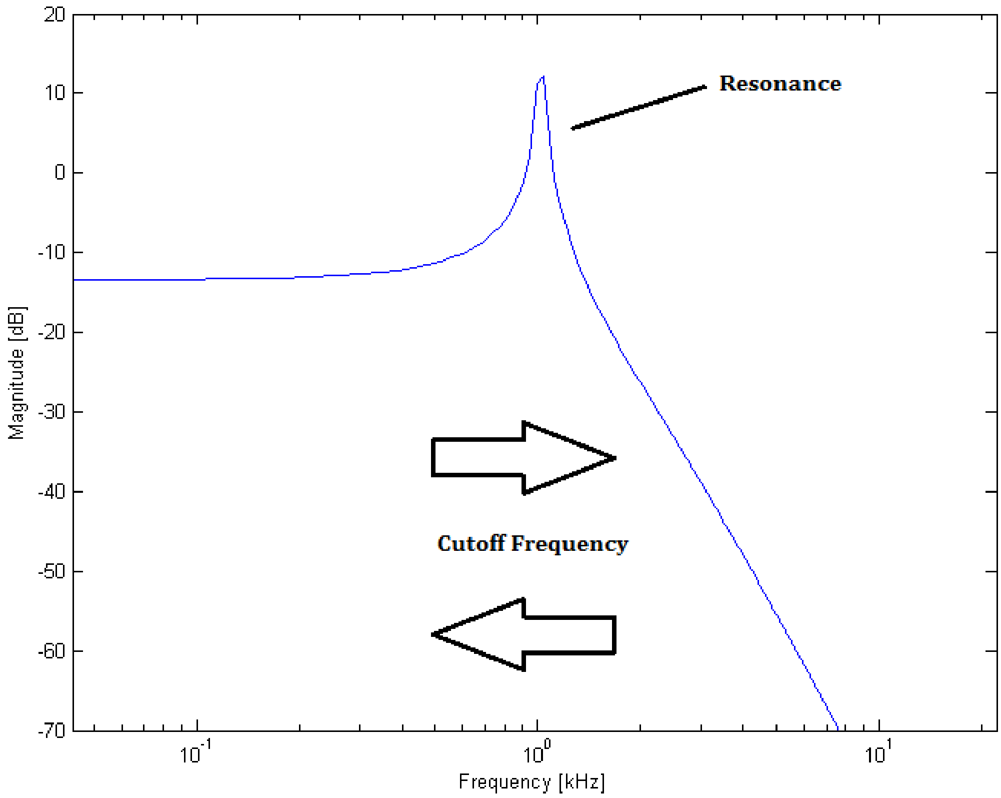

Figure 6: Variable cutoff frequency of the lowpass-type Moog filter.

2.3.4 State Variable Filter

The state variable filter is a filter which provides three different outputs: low-pass, high-pass, and band-pass. In addition to this, it allows for the independent control of the cutoff frequency and damping factor. These features prove the filter suitable for many musical applications. The difference equations of the three different outputs are given in [18] as

$$\begin{aligned} y_l(n) &= F_1 y_b(n) + y_l(n-1) \\ y_b(n) &= F_1 y_h(n) + y_b(n-1) \\ y_h(n) &= x(n) - y_l(n-1) - Q_1 y_b(n-1), \end{aligned}$$

where y_l is the low-pass output, y_b the band-pass output, and y_h the high-pass output. The tuning coefficients F_1 and Q_1 are given as

$$F_1 = 2\sin\left(\pi f_c/f_s\right) \tag{24}$$

$$Q_1 = 2\zeta. \tag{25}$$

where, f_c and ζ are the tuning parameters and f_s is the sampling frequency. In this project the filter will be used as a variable bandpass filter, which can also be used to create a wah-wah effect commonly used in guitar effects.

2.4 Spectral Subtraction

2.4.1 Basic Principle

Spectral subtraction is a method used for removing noise from a signal, cf. [19]. It is an invaluable tool used in this project enabling the use of high quality sampling synthesis. The idea behind spectral subtraction is to subtract the magnitude spectrum of the noise from the noisy signal magnitude spectrum, leaving behind the spectrum of the clean part of the signal. The noisy signal magnitude spectrum is calculated from a segment of the signal with only noise present. The spectrum of the obtained clean signal can than be converted back to the time domain to generate a noiseless audio signal. The steps behind this method will be explained in more detail below.

The noisy input signal is analyzed by taking half-overlapped windowed segments of the signal. The noisy signal in the time domain is given as the sum of the clean signal $s(k)$ and the noise $n(k)$

$$x(k) = s(k) + n(k). \tag{26}$$

Converting to the frequency domain using the Fourier transform gives

$$X(e^{j\omega}) = S(e^{j\omega}) + N(e^{j\omega}), \tag{27}$$

where j is the imaginary unit and w angular frequency defined as

$$\omega = 2\pi f \tag{28}$$

In practice, Short-Time Fourier Transform (STFT) is used to calculate the magnitude spectra of sections of the signal as it changes over time, cf. [20].

The spectral subtraction estimator

$$\hat{S}(e^{j\omega}) = \left[|X(e^{j\omega})| - \mu(e^{j\omega})\right] e^{j\theta_x(e^{j\omega})} \tag{29}$$

is obtained by replacing the magnitude $|N(e^{j\omega})|$ of $N(e^{j\omega})$ with its average value $\mu(e^{j\omega})$ (taken from only-noise part of signal) and the phase $\theta_N(e^{j\omega})$ of $N(e^{j\omega})$ with the phase $\theta_x(e^{j\omega})$ of $X(e^{j\omega})$. This estimate causes a spectral error given by

$$\epsilon(e^{j\omega}) = \hat{S}(e^{j\omega}) - S(e^{j\omega}) = N(e^{j\omega}) - \mu(e^{j\omega})e^{j\theta_x}. \tag{30}$$

2.4.2 Reducing Spectral Error

This spectral error causes unwanted audible changes in the signal and these effects should be reduced in some manner. There are four different methods available for reducing the effects of this error: magnitude averaging, half-wave rectification, residual noise reduction, and additional signal attenuation during noise-only segments. The spectral subtraction algorithm in this project will implement magnitude averaging and residual noise reduction.

Magnitude averaging uses averaging of spectral magnitudes to reduce the spectral error. $|X(e^{j\omega})|$ is replaced with $\overline{|X(e^{j\omega})|}$ where

$$\overline{|X(e^{j\omega})|} = \frac{1}{M}\sum_{i=0}^{M-1} |X_i(e^{j\omega})| \tag{31}$$

and $X_i(e^{j\omega})$ is the *ith* time-windowed transform of $x(k)$. This gives

$$S_A(e^{j\omega}) = \left[\overline{|X(e^{j\omega})|} - \mu(e^{j\omega})\right] e^{j\theta_x(e^{j\omega})} \tag{32}$$

modifying the spectral error

$$\epsilon(e^{j\omega}) = S_A(e^{j\omega}) - S(e^{j\omega}) \cong \overline{|N|} - \mu. \tag{33}$$

Residual noise can be observed as narrow bands of magnitude spikes randomly placed. Transformed back into the time-domain, these random magnitude spikes will cause unwanted sound effects and they should be removed. Residual noise is reduced by

$$\begin{aligned}
|\hat{S}_i(e^{j\omega})| &= |\hat{S}_i(e^{j\omega})|, \quad \text{for} \quad |\hat{S}_i(e^{j\omega})| \geq \max|N_R(e^{j\omega})| \\
|\hat{S}_i(e^{j\omega})| &= \min\left\{|\hat{S}_j(e^{j\omega})| j = i-1, i, i+1\right\}, \text{for} \quad |\hat{S}_i(e^{j\omega})| < \max|N_R(e^{j\omega})|
\end{aligned}$$

where $\max|N_R(e^{j\omega})|$ is the maximum value of the noise residual measured during the noise-only segment of the signal. The maximum value of the noise residual is measured using STFT over several time frames.

After the modified magnitude spectrum is obtained, the signal is then coverted back to the time-domain using inverse FFT with the overlap-add method to form the new clean output signal, cf. [19].

2.5 Peak Detection

Peak detection is the process of locating the local maxima or "peaks" of a signal. A value in a signal is considered to be a peak when its amplitude is higher than that of its neighbors. Peak detection is an important tool in signal processing, as the peaks of a signal are often of interest [21]. In this project, a peak detection algorithm is used which compares each signal value to its neighbors and concludes whether it is a peak according to the following parameters:

– Minimum peak height: the minimum amplitude required to be considered a peak

– Minimum peak separation: the minimum distance between peaks, any peaks closer than the minimum distance to another peak will not be considered peaks

– Minimum height difference: the minimum difference in amplitude required between neighboring values to be considered a peak.

Often, it is required to perform filtering on a signal before peak detection can be a viable option. Filtering is used to remove noise, as a noisy signal will easily lead to false detection of peaks. A signal should not be smoothed out too greatly with filtering though, as this can cause important and valid peaks to disappear [21]. An example of peak detection is shown in Fig. 7.

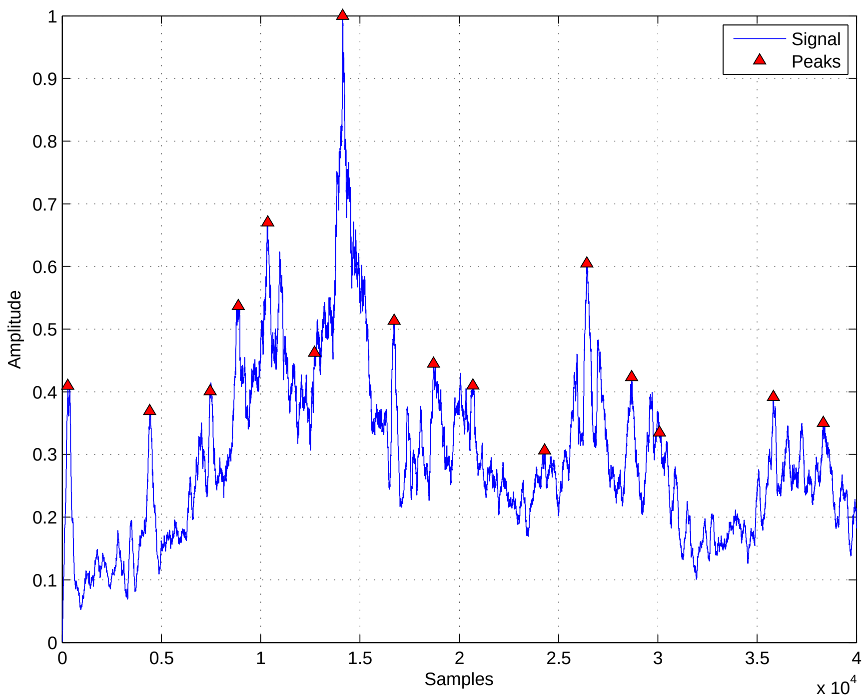

Figure 7: Peak detection used on a signal.

3 Hydraulic Sounds

3.1 Machine Hydraulics

One main component missing from the simulators was the sound of hydraulics in the machines. Hydraulics are used in operating the boom and many other features of the drill rig, harvester and forwarder and also to lift up the carriage on a truck-mounted platform. As these sounds can be quite dominant in the actual machines, it was required to come up with a way to synthesize these sounds for the simulators. However, it should be noted that in the more modern drill rigs hydraulic sounds are quite inaudible as the cabin is heavily sound proofed.

3.2 Synthesis Method

Linear predictive coding (LPC) was chosen as the most suitable method for synthesizing hydraulic sounds and the results were excellent. Using this method, synthesized WAV files were created which could then be played back in the simulators in real-time. LPC, as explained in Sec. 2.2, is used to extract the spectral envelope of the original signal, which can then be used as a basis for sound synthesis. Using LPC, we can calculate a filter corresponding to the spectrum of the original hydraulic sound signal, which is then used to filter a white noise excitation resulting in a synthesized hydraulic sound. In this case, extremely high filter orders were used ($p = 1000$), as there was no need for real-time processing and these values achieved exceptional results. The lengths of the original hydraulic sound samples, from which the spectral envelopes were extracted, ranged from 100 ms to 150 ms. An advantage to using this method is the possibility of synthesizing sound signals of arbitrary length, i.e. the length of the white noise excitation determines the length of the synthesized signal. This allows for the possibility of creating longer sounds, which can prove difficult if only sample based synthesis is used due to the periodic sound caused by looping the same short sample. In addition, by synthesizing several different samples with different white noise excitations, the signals will not cause a periodic sound even when looped.

3.3 Synthesized Hydraulic Sounds

This section will cover the different types of hydraulic sounds that were synthesized. All the original recording samples used for the basis of the synthesis are taken from a longer reference recording. This main file is a minute long recording of the hydraulic sounds in a drill rig presenting several different sound events possible in a hydraulic system.

3.3.1 Basic Hydraulic Sound

The most basic and prominent hydraulic sound is the basic "hiss" sound, which can be seen in Fig. 9 below. The LPC filter is calculated from a 100 ms portion of the original hiss sound and then a white noise excitation is filtered with it. Figures 8 and

9 present the different stages of the synthesis of this sound: the spectral envelope of the original signal calculated using LPC and the white noise excitation signal in Fig. 8 and the time and frequency domain signals of the original and synthetic signal in Fig. 9.

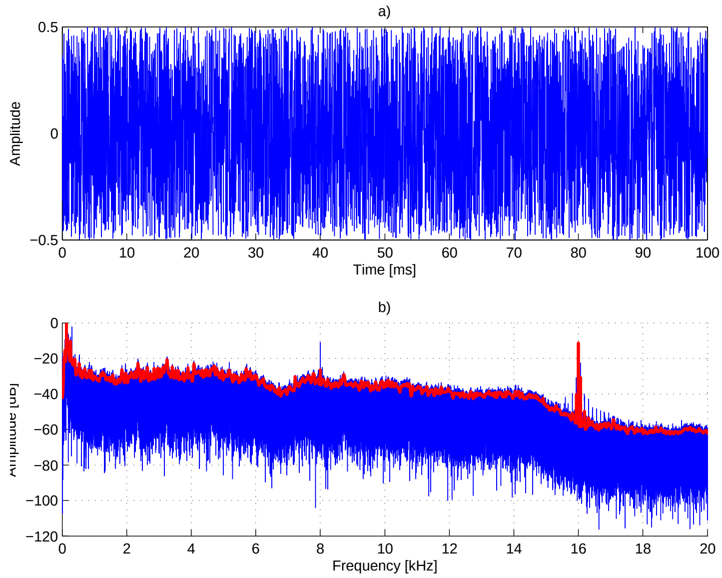

Figure 8: a) White noise excitation signal and b) spectrum of original signal in blue and LP spectrum in red.

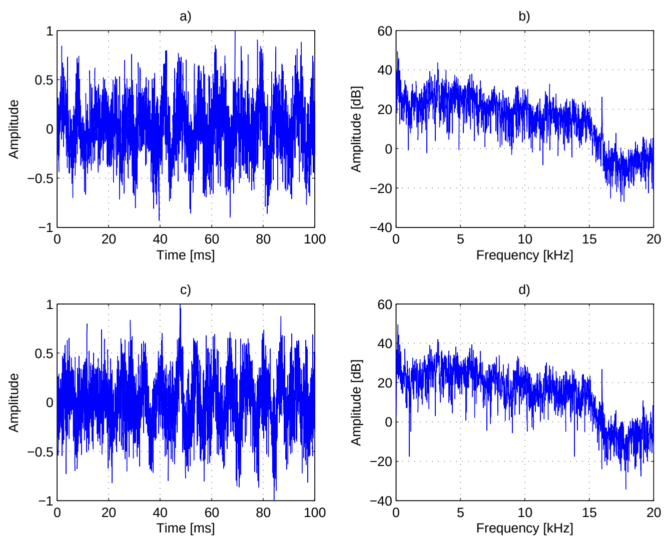

Figure 9: Time domain presentations and spectra of the a)b) original signal and c)d) synthetic signal.

3.3.2 Fading Hydraulic Sound

The fading hydraulic sound is a sound which occurs as the hydraulic fluids stop flowing within the system. This sound was synthesized first by using an LPC filter (calculated from a 150 ms portion of the original signal) with an order of $p = 1000$ and a white noise excitation to achieve the basic hydraulic sound. Figure 10 presents the spectrum of the original signal, the LP-spectrum, the white noise excitation, and the spectrum of the synthesized signal.

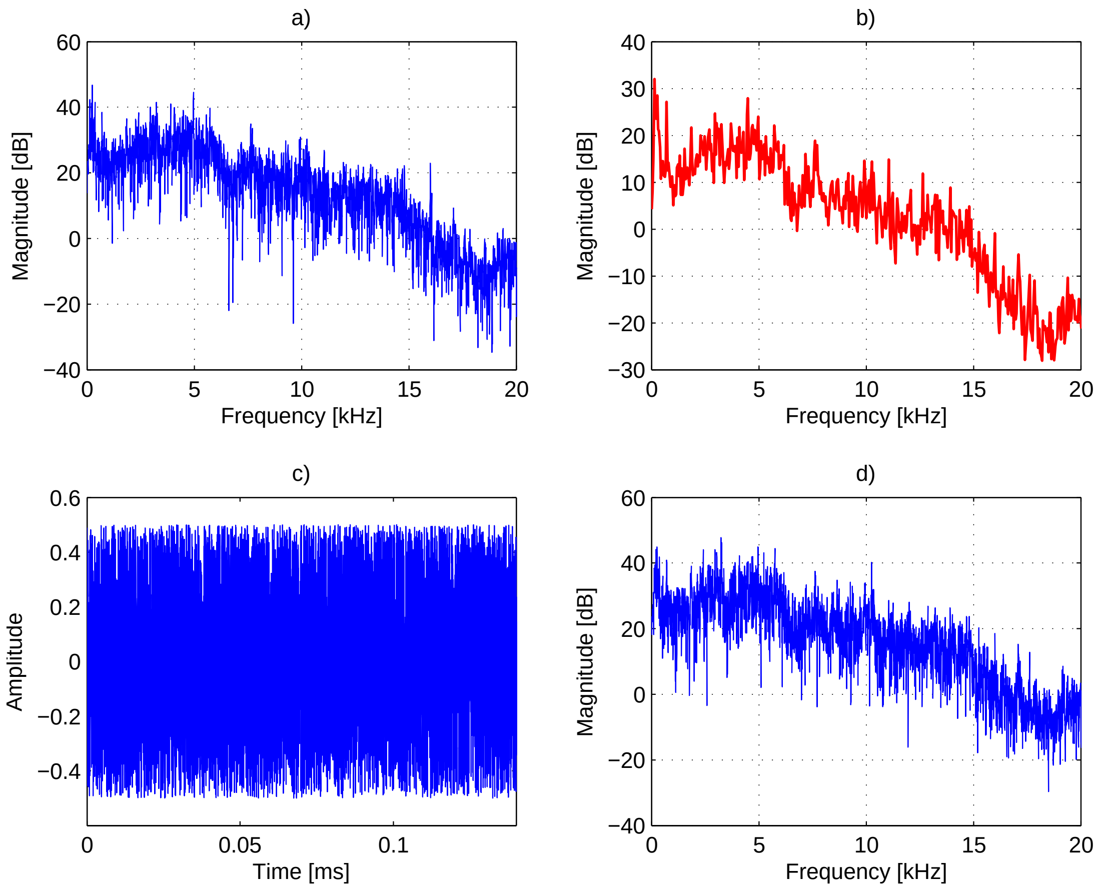

Figure 10: a) Spectrum of original signal, b) LP spectrum, c) white noise excitation, and d) spectrum of synthesized signal.

To simulate the sound of the hydraulic fluids stopping their flow, a Moog filter, explained in Sec. 2.3.3, was used to sweep a low-pass filter to a cutoff frequency of 700 Hz with a logarithmic frequency envelope. In addition, a linear amplitude envelope was used to fade out the sound level at the end of the sample. The basic hydraulic sound was multiplied with the amplitude envelope and the output of the Moog filter was added to the end. The output of the Moog filter is the basic hydraulic sound filtered by the Moog filter with a low pass filter sweep from 2200 Hz to 700 Hz. The cutoff frequency envelope, the amplitude envelope, and the output of the

Moog filter are presented in Fig. 11.

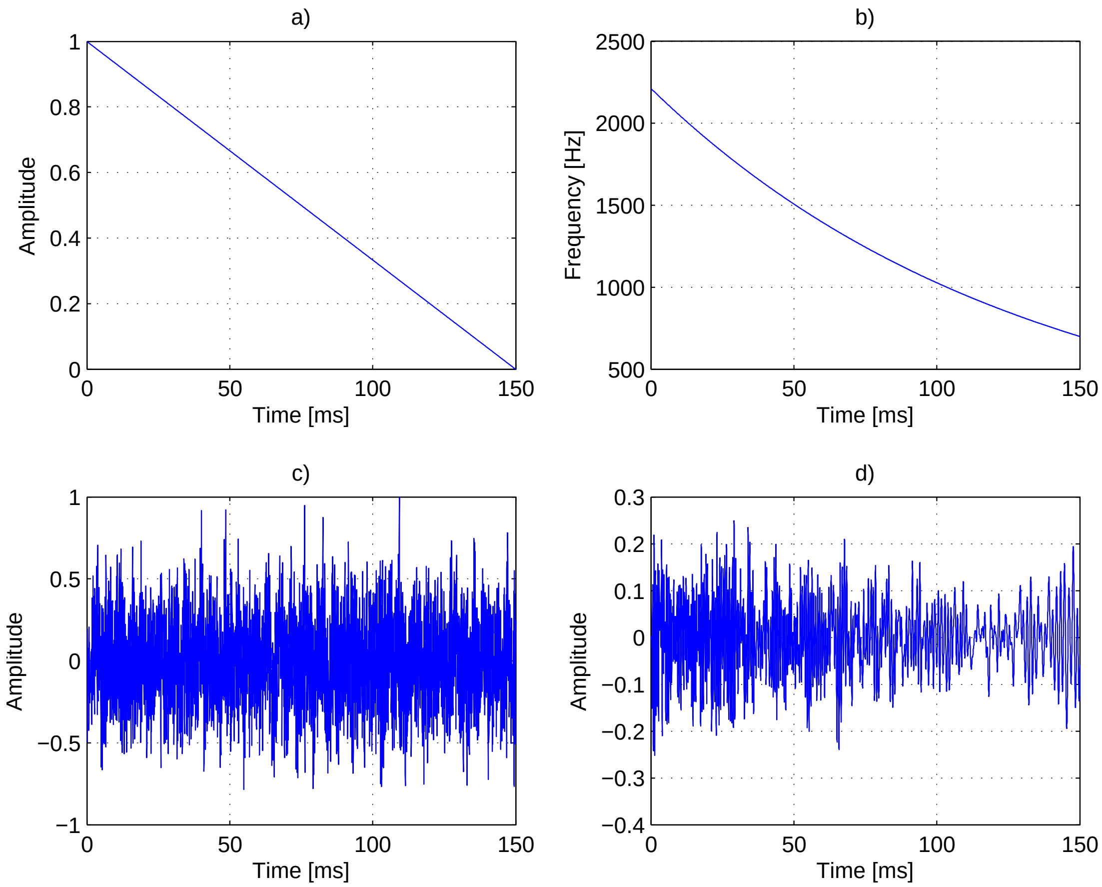

Figure 11: a) Amplitude envelope, b) cutoff frequency envelope, c) original hydraulic sound without fade out, and d) Moog filter output.

3.3.3 High Frequency Hydraulic Sound

In addition to the basic hydraulic sound, a higher frequency sound was also required. To achieve this, a high frequency hissing sound (100 ms) was first band-pass filtered from a recording of actual hydraulic sounds. The high frequency hydraulic sound could be heard in the frequency range of 4.5 to 8.5 kHz, thus a band-pass filter meeting these specifications was implemented. The band-pass filter utilized a pass band of 4.5 to 8.5 kHz and with transition bands at 4.4 to 4.5 kHz and 8.5 to 8.6 kHz. A Kaiser-window, explained in Sec. 2.3.1, was used as the windowing function for the design of the FIR digital filter, which can be seen in Fig. 12. The spectrum of the original and filtered signal is presented in Fig. 13.

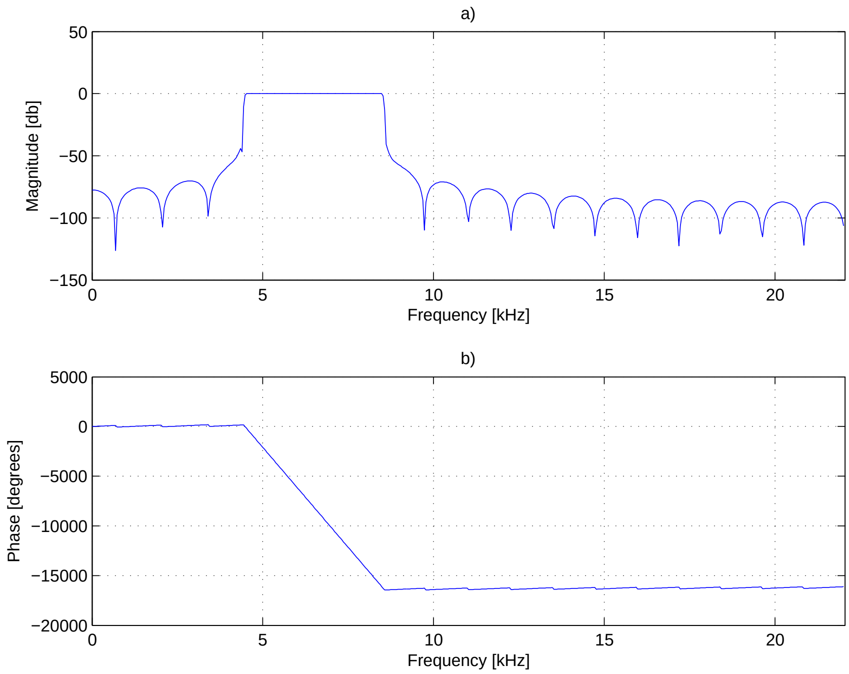

Figure 12: a) Frequency response and b) phase response of the band-pass filter with pass band at $[4.5, 8.5]$ kHz.

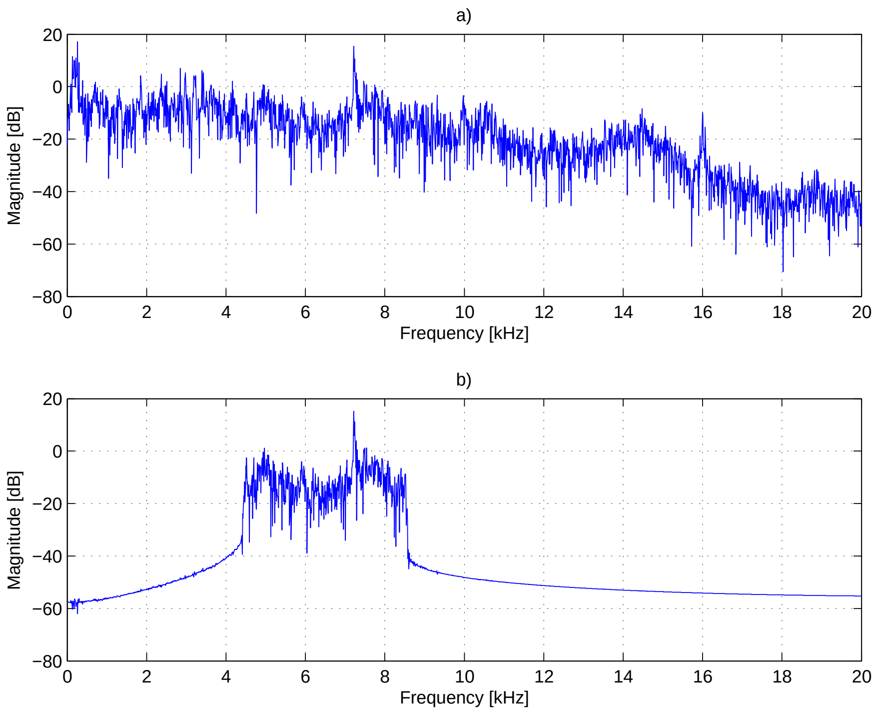

Figure 13: Spectrum of a) the original signal and b) band-passed signal.

After obtaining the filtered output, linear predictive coding could again be used to extract the spectral envelope of the required sound sample. LPC was used in the exact same manner as with the basic hydraulic sound, i.e. filter order $p = 1000$ and white noise as the excitation signal. Figure 14 presents the filtered spectrum, the LP spectrum, the white noise excitation, and the spectrum of the synthesized signal.

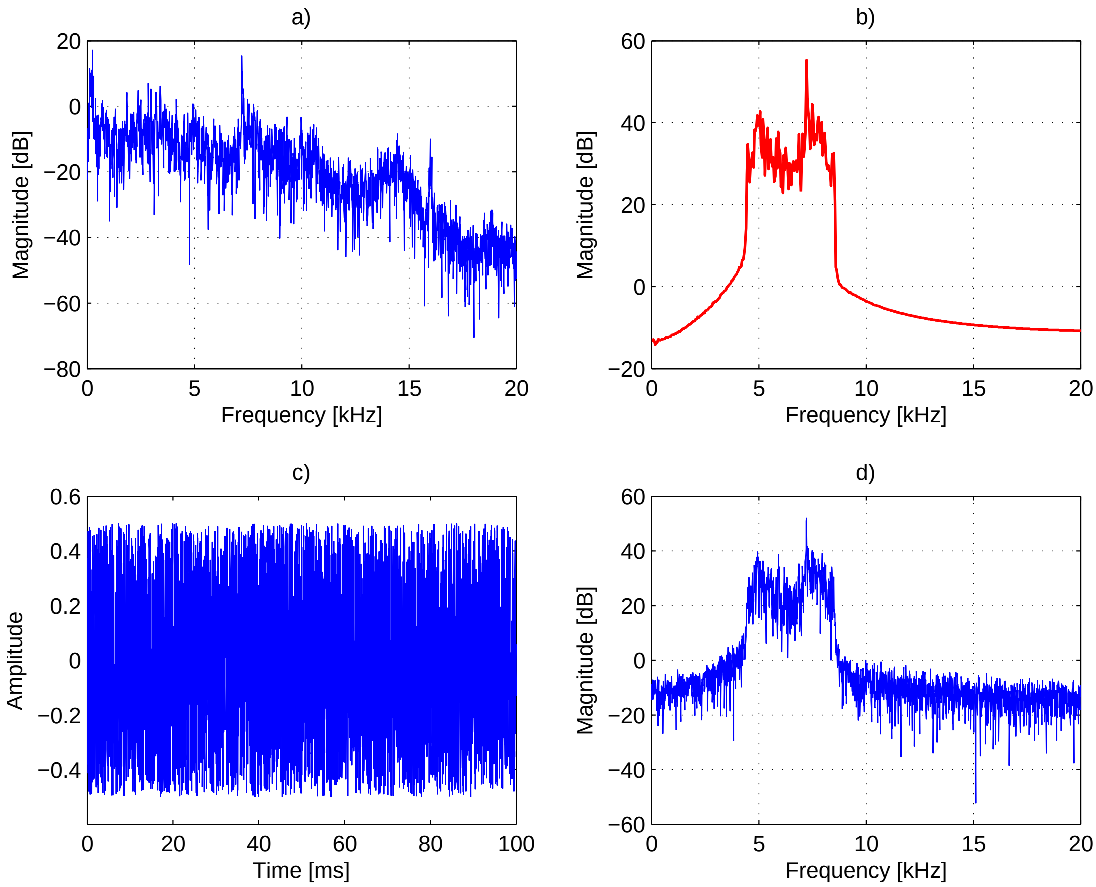

Figure 14: a) Spectrum of original signal, b) LP spectrum, c) white noise excitation, and d) spectrum of synthesized signal.

3.3.4 High Frequency Variable Hydraulic Sound

The high frequency variable hydraulic sound is a synthesized sound used to try and imitate the sound of the hydraulic fluids trying to squeeze through the system. The sound is a high frequency sound which varies in frequency, causing a squealing like sound. The sound is synthesized from the synthesized output of the high frequency hydraulic sound presented in the previous section, by further filtering it with a state variable filter, which is commonly known for the "wah wah" effect used in guitar effects. This filter alters the band-pass of the filter higher and lower in frequency, causing a high frequency "wah wah" effect as seen in the spectrogram in Fig. 15.

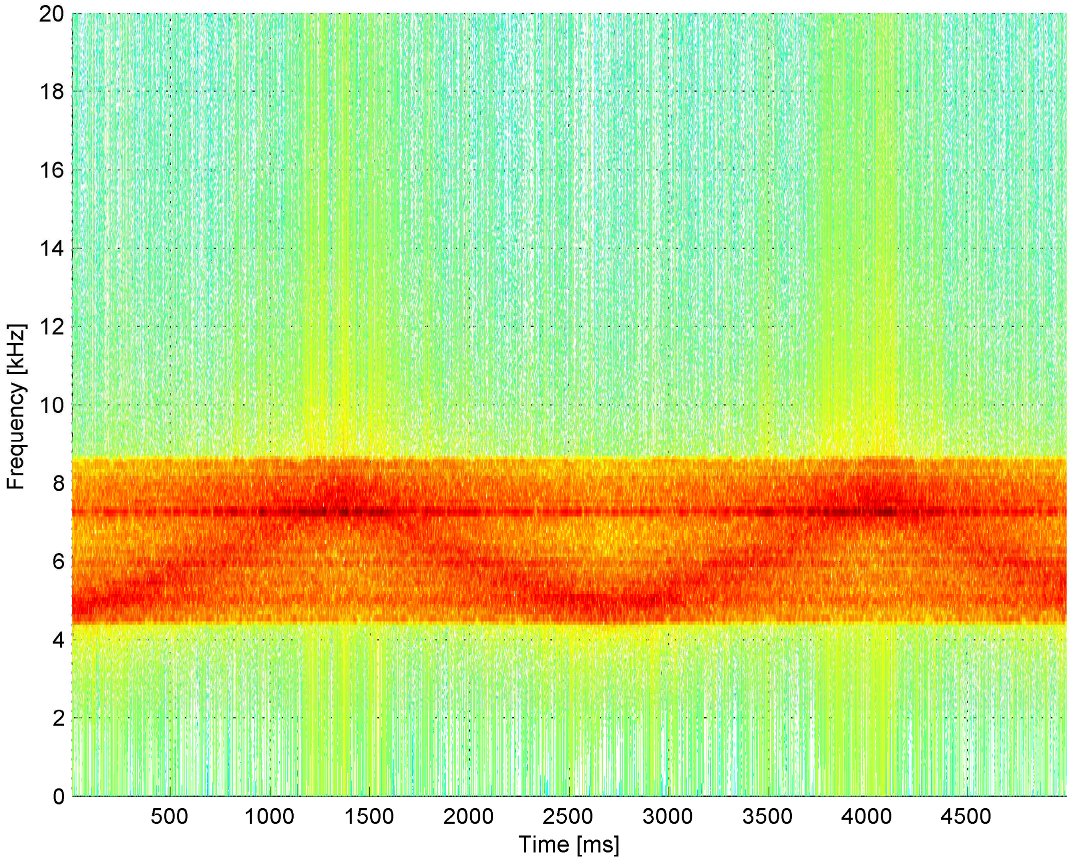

Figure 15: Spectrogram of the high frequency variable sound.

3.3.5 Hydraulic Piston Contact Sound

The hydraulic system also consists of contact sounds or "thump" sounds caused by the pistons of the machine. These sounds were also synthesized using linear predictive coding with a filter order of $p = 1000$, but instead of a white noise excitation, a simple impulse was utilized. Using LPC to create synthetic versions of the sounds allowed for noiseless samples, as the original signals contained some background noise. Several different LPC-filters were calculated from different contact sound samples from the recordings and the same impulse excitation was filtered with all filters. The impulse excitation and the spectra of the LPC filters can be seen in Fig. 16. Each synthesized sample differed from each other due to the different LPC filters, allowing for a varying pallet of contact sound samples. The time and frequency domain presentations of these synthesized samples can be seen in Figs. 17, 18, and 19. The original and synthetic signals differ slightly in the time-domain, but perceptually the signals sound very similar, which can be noticed in the similarity between the original and synthetic spectra.

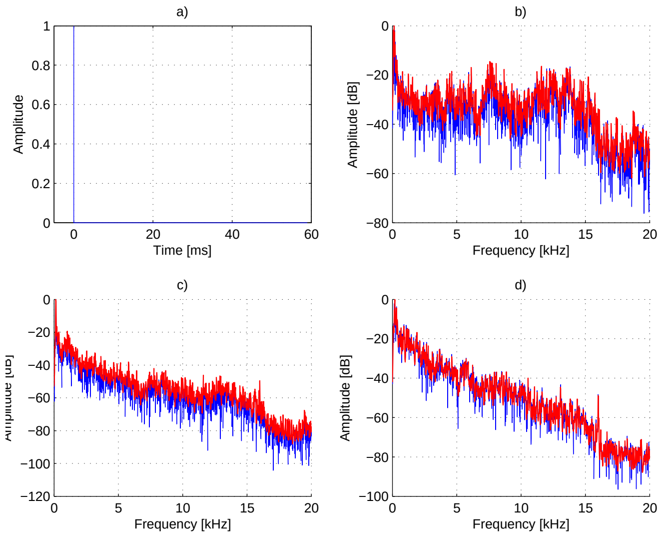

Figure 16: Impulse excitation in a) and the original spectrums (blue) and LP spectrums (red) of b) thump 1 c) thump 2 d) thump 3.

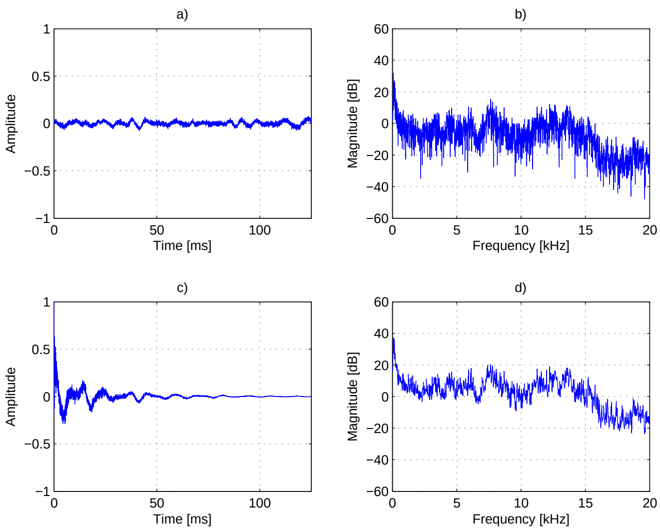

Figure 17: Thump 1: The original signal presented in a) the time domain and b) frequency domain. The synthesized signal presented in c) the time domain and d) frequency domain.

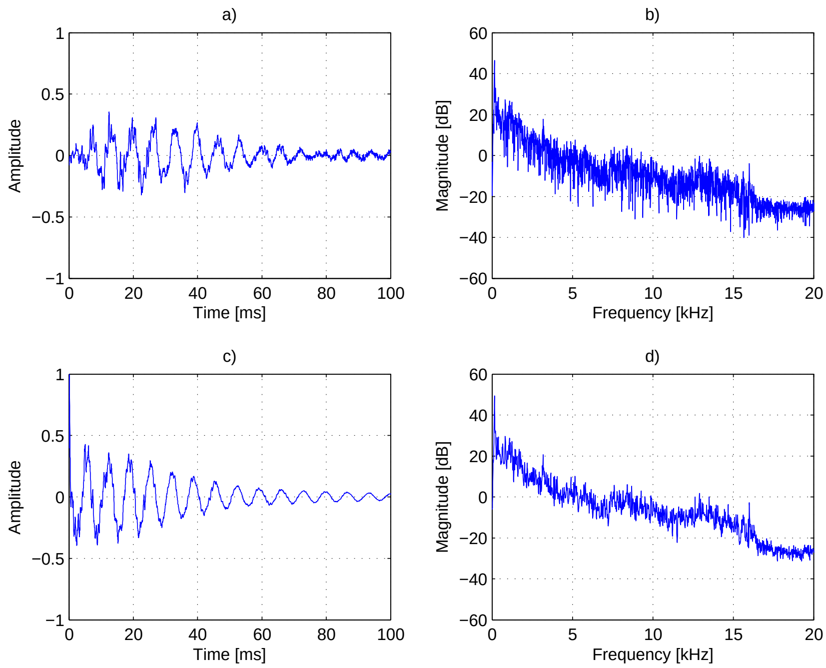

Figure 18: Thump 2: The original signal presented in a) the time domain and b) frequency domain. The synthesized signal presented in c) the time domain and d) frequency domain.

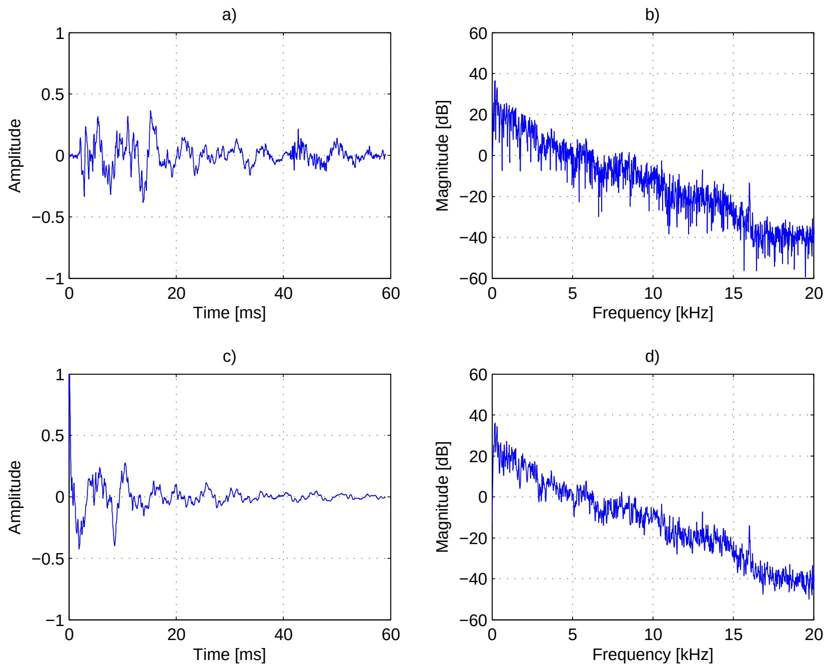

Figure 19: Thump 3: The original signal presented in a) the time domain and b) frequency domain. The synthesized signal presented in c) the time domain and d) frequency domain.

As can be seen in Figs. 17, 18, and 19 above, most of the energy in the spectra is concentrated around the lower frequencies. The spectra of the original and synthetic contact sounds are presented between 0 to 4 kHz in Fig. 20 below, allowing for a more detailed comparison of the spectral characteristics.

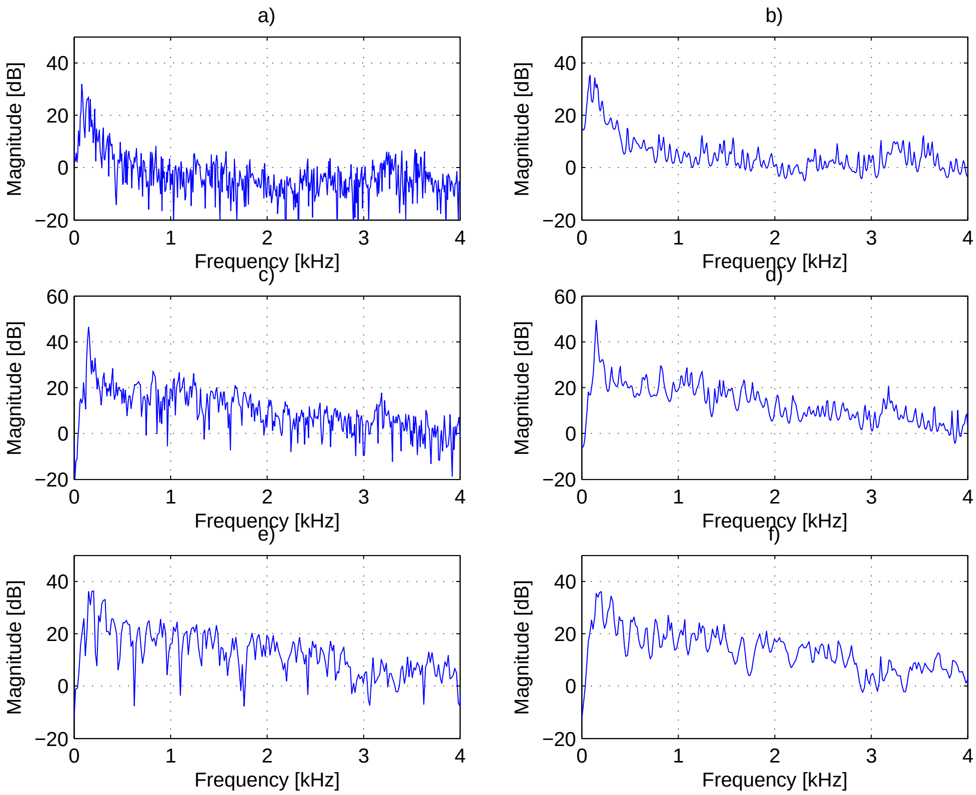

Figure 20: Spectra of a) thump 1 original signal, b) thump 1 synthetic signal, c) thump 2 original signal, d) thump 2 synthetic signal, e) thump 3 original signal, and f) thump 3 synthetic signal.

4 Drilling Sounds

Drilling sounds are obviously an essential part of any drill rig simulator, as the soundscape of a real life drill rig in operation mostly consists of the sound of the rock drill. The main component of the rock drill is the piston, which converts hydraulic energy into mechanical energy to actually cause a drilling motion in the system. The boom of a drill rig holds several extra rods, which are connected together to form drill strings connected to the rock drill allowing it to dig deep into the ground [22]. Although this thesis will not go into more detail about the specifications and types of rock drills and drill rigs, it is worth mentioning that the drilling sounds analyzed in this project are from percussive hydraulic rock drills. More information about rock drills can be found in [22].

The rock drill and its rods cause five main sound types typically found in drilling situations:

1. Normal drilling: the typical drilling sound which should be heard when drilling.

2. Underfeed: feeding is the method of pushing the drilling tool against the rock. In an underfeed situation, the drill is not pushed hard enough against the rock and the sound can be easily distinguished from normal drilling by experienced drill rig operators.

3. Overfeed: overfeeding is caused by pushing the drill too hard against the rock. This causes a slight variation in sound compared to normal drilling, but compared to underfeeding, it is not as easily recognizable even to experienced operators.

4. Rattling (threads closed): rattling is the sound of the rods being removed from the rod string. The rods are shook heavily to open the threads holding the rods together. In this situation the threads are still closed causing a clearly distinguishable sound.

5. Rattling (threads open): the sound of rattling when the threads finally open is a very short but extremely loud and high frequency sound event, which is clearly different compared to rattling with the threads still closed.

The aim in this project was to first synthesize a normal drilling sound and then by utilizing statistical analysis of different drilling situations, change the parameters of the normal drilling sound to synthesize the other drilling situations.

4.1 Synthesis Method

The previous drilling sound of the drill rig simulator was used as the basis for the new synthesis method. The original drilling sound was created by simply looping a sound file of four single synthetic strikes from a rock drill, cf. Fig. 21, which will henceforth be referred to as 'clicks'. The drilling frequency was artificially altered using pitch shifting, which causes the sound to vary in frequency, but the actual frequency between the clicks does not change. Optimally, as in real-life, increasing the drilling frequency should cause the time between clicks to decrease.

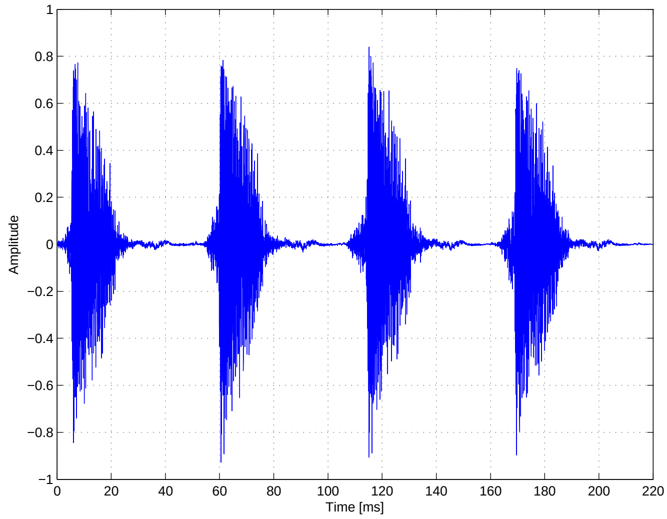

Figure 21: Time domain presentation of the previous drilling sound sample.

To create a synthetic drilling sound which varies the distance between clicks according to the drilling frequency, single clicks had to be first extracted from the original sound sample as seen in Fig. 22. The next step was to simply create a function which repeats a single click at the required frequency. By doing this, drilling sound samples of any given frequency could be correctly synthesized by just changing the distance between single click sounds.

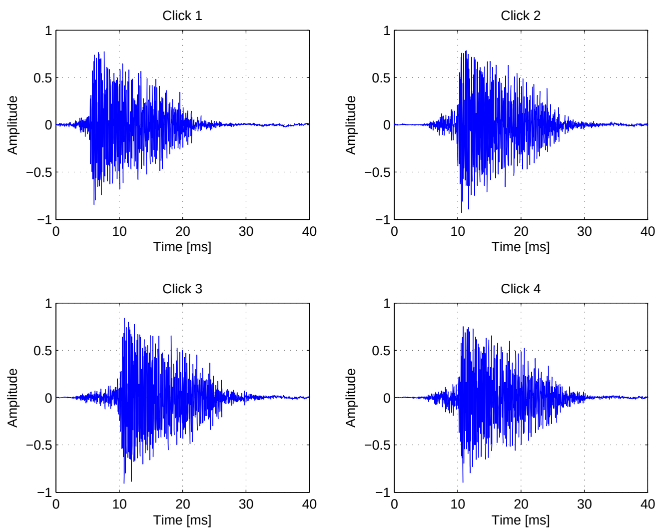

Figure 22: The four diffecent click sounds in the time domain.

To add variance in the drilling sounds, the single clicks were chosen at random each time from the four possible click samples. Although this is a viable method of adding realism to the sound, in this case it proved ineffective as the four different click sounds were too similar to each other. This can be seen from their almost identical spectra in Fig. 23 and time domain presentations in Fig. 22.

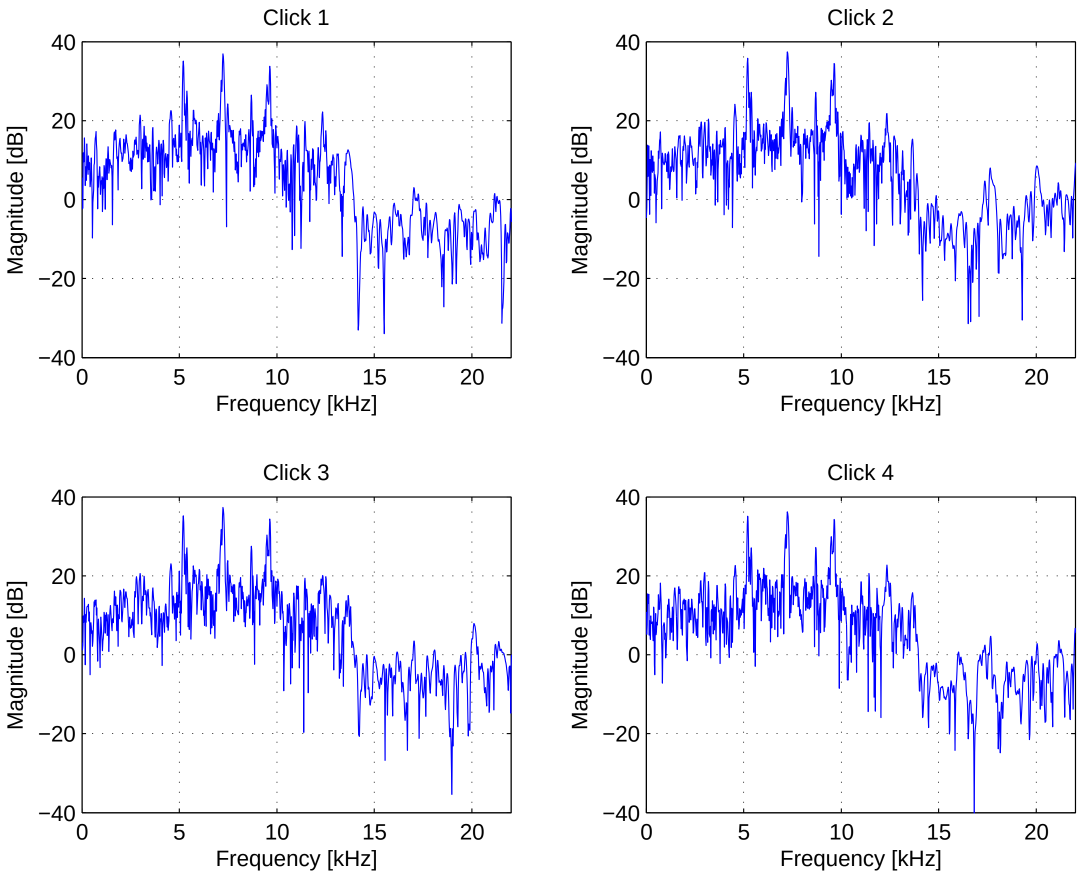

Figure 23: The spectrums of the four different click sounds.

4.2 Statistical Analysis

Although the synthetic drilling sound explained above now correctly simulates the actual drilling frequency, the resulting sound is still quite synthetic and monotonous. This is due to the clicks being almost identical, and even more so due to the fact that the distance between clicks is exactly the same at a given frequency. In actual drilling the amplitude of the clicks and the distance between them varies slightly, as each strike of rock is slightly different every time. In normal drilling the variations are quite minimal, but with, e.g. an underfeeding situation the variations can be quite drastic as the rock drill is not in firm contact with the rock. Statistical analysis of the different drilling situations is used to simulate the affects of these variations in the hopes of achieving a more realistic sound.

Actual drilling sounds of normal drilling, underfeeding, overfeeding, and rattling were analyzed to obtain values for two main parameters:

1. Amplitude variation: how much the amplitude of a single click differs from the previous click.

2. Distance variation: the amount of samples between two neighboring clicks.

A more realistic drilling sound with actual real-world data could be achieved by implementing these obtained values in the synthetic drilling sound.

Statistical analysis of the aforementioned drilling sounds was accomplished utilizing full-wave rectification, a sliding average filter, and peak detection. First, a sound sample of actual drilling was full-wave rectified and a sliding average filter was used to extract the amplitude envelope of the signal, cf. Sec. 2.3.2. From this smoothed signal, the peaks were detected using a peak detection algorithm, cf. Sec. 2.5. A single peak is assumed to correspond to the energy maximum of a single click, and thus by calculating the amplitude and distance variations between these peaks, the required data for the two parameters could be obtained.

The extracted amplitude envelopes of a normal drilling sound, underfeed sound, and overfeed sound are presented in Fig. 24. As is evident from the red line in the figure, the sliding average filter aims to smooth out the signal leaving only the main peaks and valleys visible. An optimal envelope should travel through the local maximums, but not drop drastically at the minimum points. In this case a width of 2000 samples (40 ms) seemed to work fairly well for the filter.

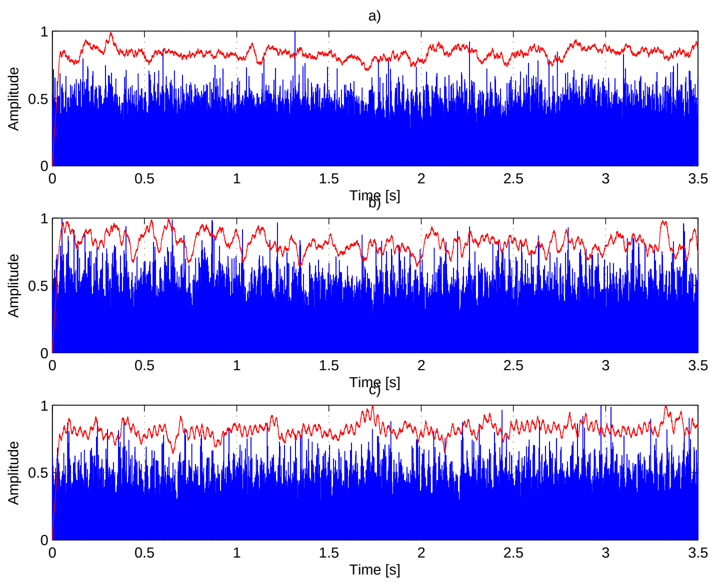

Figure 24: Full-wave rectified signals (blue) and amplitude envelopes (red) of a a) normal, b) underfeed, and c) overfeed drilling sound.

Peak detection was implemented on the previously obtained amplitude envelopes with a minimum peak height of 0.7, minimum peak seperation at 1000 samples, and minimum height difference set to zero. Figure 25 shows 500 ms segments of these signals with the peaks indicated with red triangles.

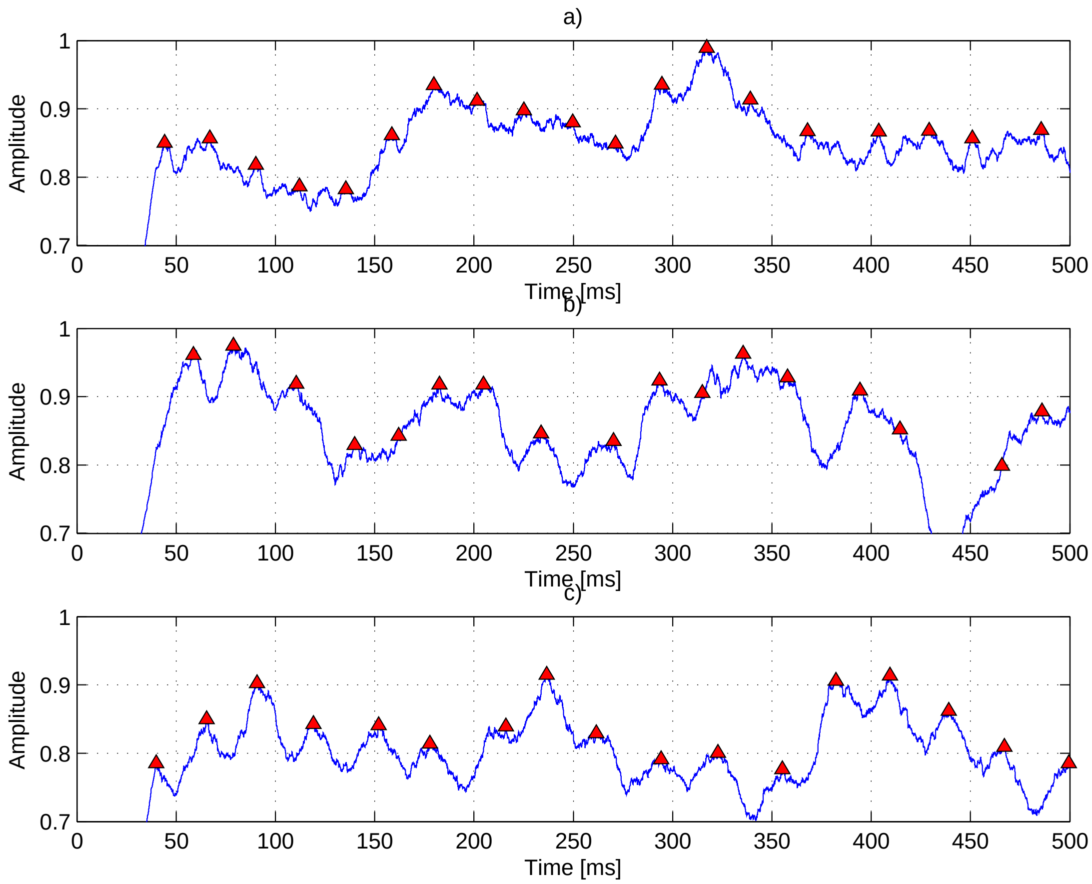

Figure 25: The detected peaks (red) of a a) normal, b) underfeed, and c) overfeed drilling sound.

Amplitude and distance variation data was calculated using these peaks and the values were implemented in the synthesis of the drilling sounds. Figure 26 presents histograms of the obtained data. All the histograms were expected to follow the shape of a normal distribution, but as is evident from the figure below, only the overfeed sound follows a Gaussian distribution in both the amplitude and distance variations. All three sounds do follow a normal distribution in relation to the amplitude variations, but for the normal and underfeeding sound the distance variations follow a distribution more closely resembling a Chi-squared distribution [23]. Noise in the original sound samples might be a cause of this phenomenon, as peak detection might fail if some peaks are mistaken as energy maxima (clicks), when they should not. This would cause the distance data to be faulty leading to unexpected results.

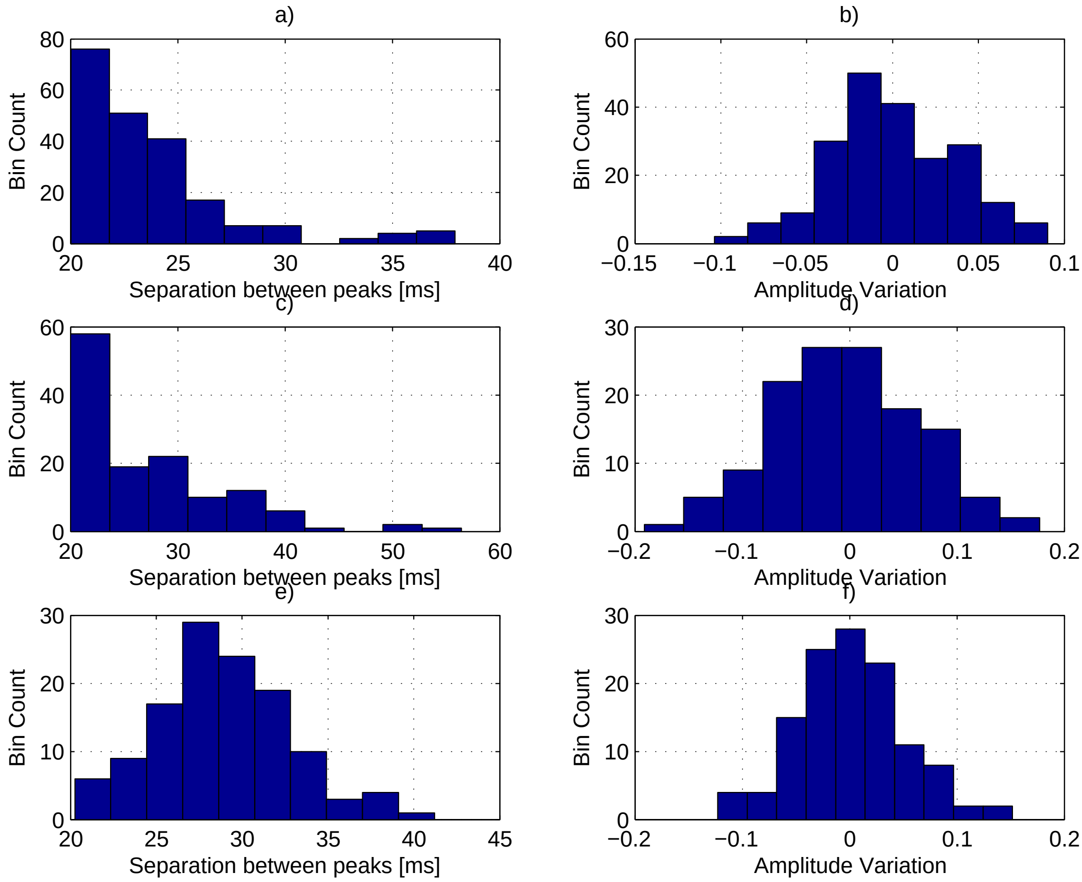

Figure 26: Histograms of a)b) normal, c)d) underfeed, and e)f) overfeed drilling.

4.3 Conclusions and Improvements

Although the proposed method of adding variance into the synthetic sound signals does infact add some realism to the overall sound (especially in differentiating between different drilling types), the sounds can still quite clearly be distinguished as synthetic. This is mostly due to the original click sounds used. No matter how realistically the distances and amplitudes of the clicks are varied, they will still sound synthetic as the source itself is synthetic sounding. An improvement to this would be to obtain noiseless real-life sound samples of single clicks, which could then be implemented in the same fashion as the synthetic ones. Separate sound samples for normal drilling, underfeeding, overfeeding, and the two types of rattling would yield a much higher quality synthetic result. Rattling sound synthesis would especially benefit from this as the click sounds in rattling are very different compared to drilling. The absence of a rattling sound click is also one of the reasons why rattling sounds were not synthesized in this part of the project.

5 Harvester and Forwarder Sounds

One main focus of this project was the synthesis of typical sounds found in two important forest machines: the harvester and forwarder. Among the required sounds for the simulator for these machines were: feeding, delimbing, several different types of contact sounds, and hydraulic sounds which were already explained in Sec. 3.

5.1 The Machines

The harvester and forwarder are typically used together in forestry: the harvester is used to cut down trees and delimb them and the forwarder is used to transport the logs from the forest. In addition to the increased speed and effectiveness these machines bring to logging, they also provide a safer work environment as the operators are stationed safely inside a protected driving cabin, away from the dangers of falling trees and motorsaws.

A modern forest harvester can be seen in Fig. 27. The machine typically consists of a cabin, a diesel engine, wheels or tracks, an extendable boom, and a harvester head. The diesel engine is used to power the vehicle and the boom combined with the harvester head through a hydraulic system. The combination of a powerful diesel engine and heavy duty wheels or tracks make forest machines mobile and robust on terrains of all kinds. The boom is used to move around the harvester head, which consists of the tools required for felling and delimbing trees. The average harvester head employs a chain saw for cutting the tree, delimbing knives used for removing branches, and feed rollers for moving the tree through the harvester head. The last two parts of the harvester head are mostly of interest in this project as they are the cause of two important sound events: feeding and delimbing.

Figure 27: A forest harvester. Retreived from www.ponsse.com

A forest forwarder is presented in Fig. 28. Similar to the harvester, it also consists of a cabin, diesel engine, wheels or tracks, and an extendable boom. Instead of a harvester head, the forwarder employs a grapple at the end of the boom. The grapple is used to grab various amounts of logs and move them on and off the carriage. The grapple also contains a load brake, which is used to prevent the grapple from swinging too drastically. The carriage is used to store and transport the logs on the machine. The carriage consists of the bunks (vertical poles at the side) and the screen (the back of the carriage). The grapple, logs, bunks, screen, load brake, and hydraulics are the cause of the sounds which will be analyzed later below.

Figure 28: A forest forwarder. Retreived from www.ponsse.com.

5.2 Feeding Sound

Feeding is the process of moving a tree trunk through the harvester head. Moving the log allows for it to be sawed at desired positions and also to allow the delimbing knives to remove branches from the log. The feeding sound is caused by feeding a section of the log with no branches through the harvester head. The sound is the noise of the feed rollers scraping against the bark of the tree.

A video of a forest harvester at work was used as the starting point for the synthesis of a feeding sound. A segment of audio containing a feeding event was extracted from the video and analyzed and processed for the final synthesis steps. The extracted sound sample contained a fair amount of background noise, including engine and wind noise, and thus spectral subtraction, presented in Sec. 2.4, was required to achieve a cleaner sample. Even after spectral subtraction, some engine noise was still present in the sample which would adversely affect the quality of the synthesis. To attenuate the effect of the engine noise, two FIR filters (Sec. 2.3.1) were implemented:

1. A low-pass filter with a cutoff frequency at 8 kHz and a filter order of 99
2. A band-stop filter (notch filter) with a stopband at 500-600 Hz and a filter order of 1970

The low-pass filter was used to remove some of the high frequency components caused by the diesel engine of the harvester. Next, the notch filter was implemented to remove a very distinct and loud squealing sound between 500-600 Hz. The effects of filtering the original sound sample with these two FIR filters can be seen in the spectrograms shown in Fig. 29.

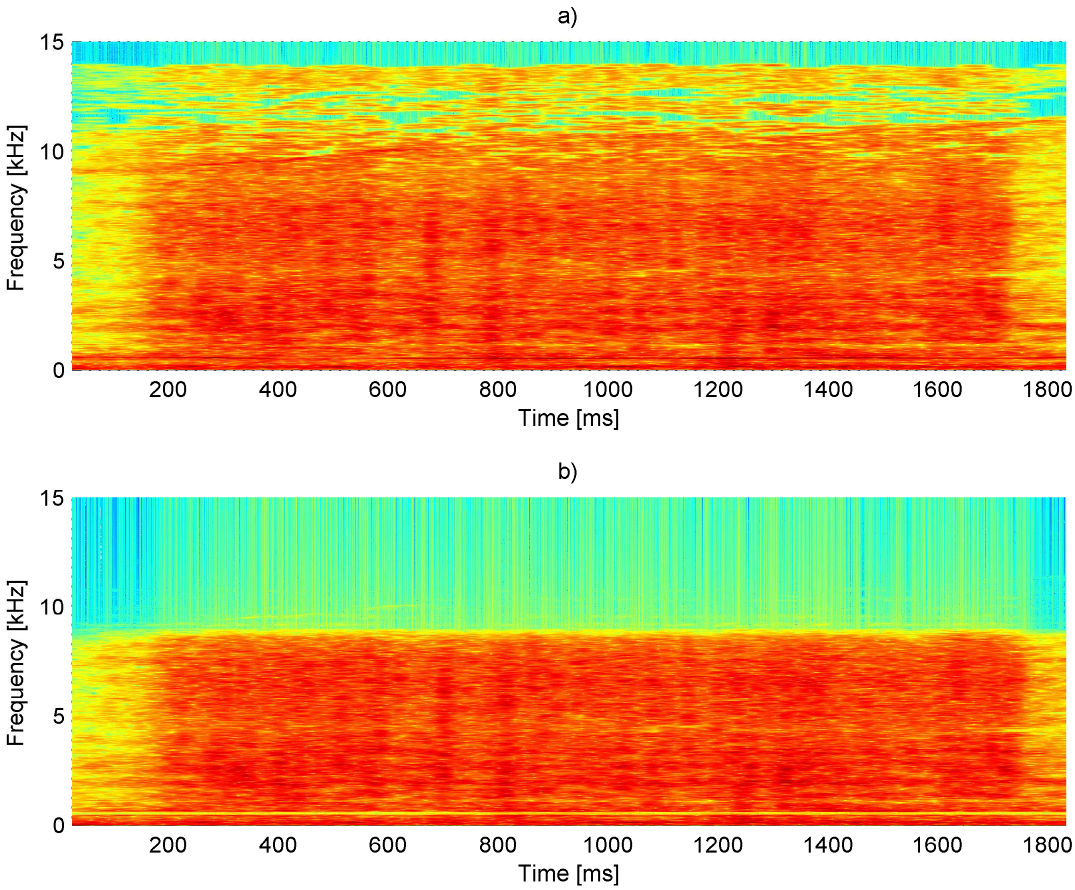

Figure 29: Spectrograms of a) original signal (spectral subtraction implemented) b) filtered signal.

By implementing spectral subtraction and filtering with the two aforementioned FIR-filters, the sound sample was clean enough to be used as a reference sample for the synthesis. An LPC filter was calculated from the filtered sound sample with a filter order of $p = 1000$. As explained in Sec. 2.2, the LPC filter contains the spectral characteristics of the original signal. A white noise excitation signal was filtered with the LPC filter leading to a purely synthetic signal with the same spectral envelope as the original signal.

The next step in the synthesis was to correctly model the amplitude envelope of the original signal which is presented in Fig. 30. This was achieved with full wave rectification and a sliding average filter using the method explained in Sec. 2.3.2. A window size of $N = 200$ was used for the sliding average filter. The achieved amplitude envelope was then multiplied with the output of the LPC synthesis and as a result, a purely synthetic signal was obtained with the correct amplitude and frequency characteristics.

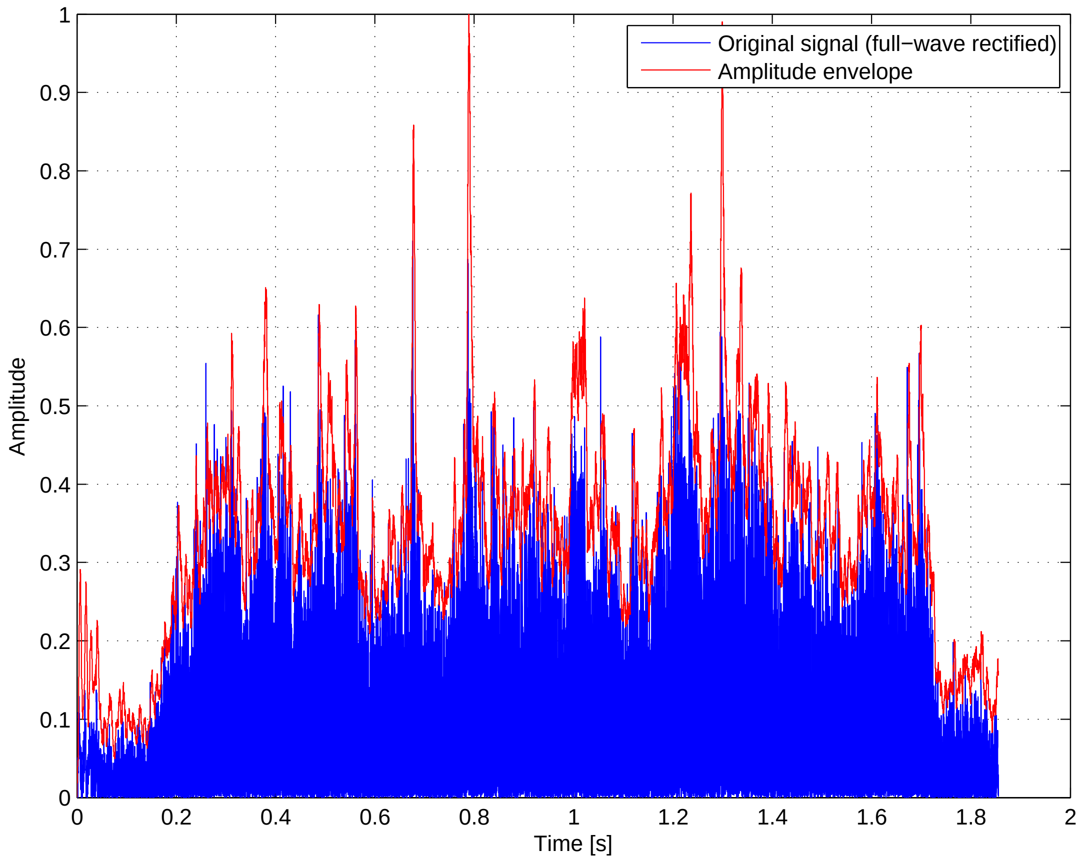

Figure 30: Time-domain presentation of the full-wave rectified original signal and the amplitude envelope calculated using a sliding average filter.

The time domain presentations and spectrograms of the original and synthetic signals can be observed in Fig. 31. Both signals look fairly similar in both the time domain and frequency domain, implying a successful synthetic result.

5.3 Delimbing Sound

The delimbing sound is closely related to the feeding sound presented above. The sound is caused by branches being cut by the delimbing knives in the harvester head as the log is being fed through it. So essentially, there will also always be a feeding

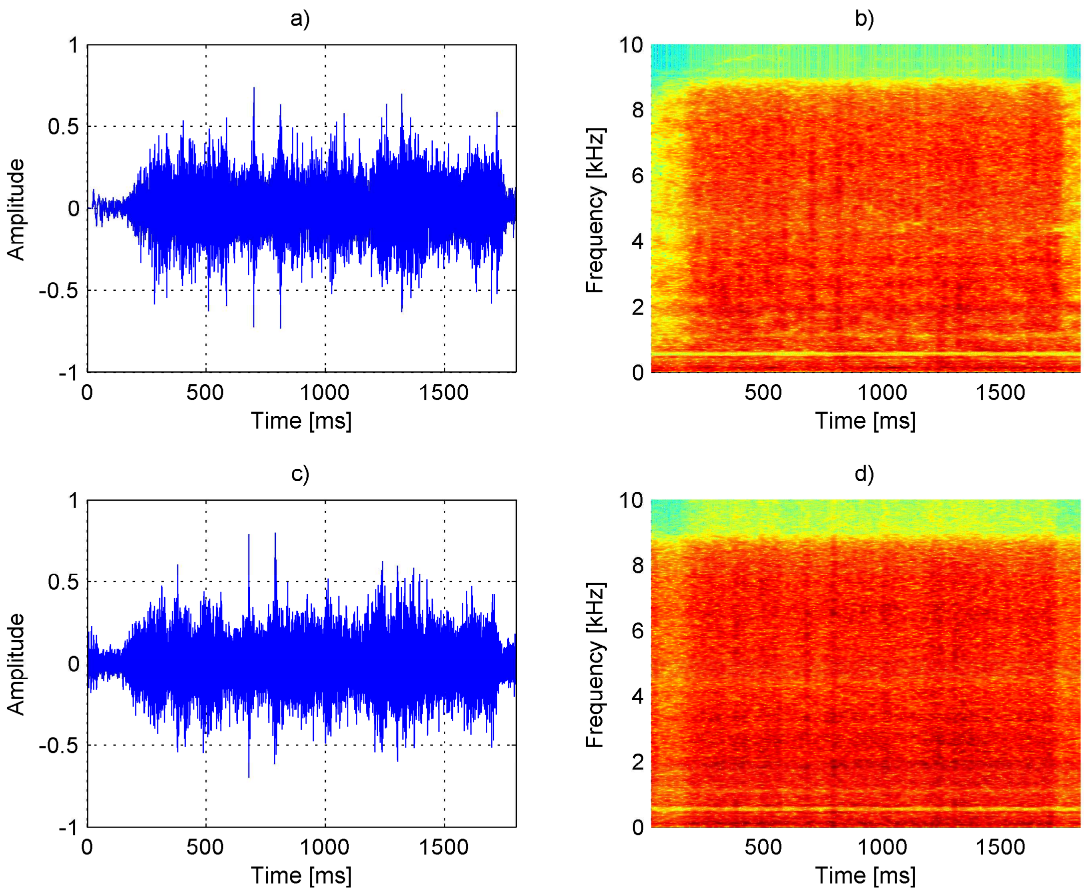

Figure 31: Time domain presentations and spectrograms of the a)b) original and c)d) synthetic feeding sound.

sound underneath the delimbing sound, as feeding is required for delimbing to take place. The approach chosen to synthesize the sound of a delimbing event was to create audio samples of branches being cut and combine them with the previously synthesized feeding sound. This allows for two separate types of sounds which can then be easily combined or played separately depending on the situation as the log is being fed through the harvester head:

1. No branches on the log $\Rightarrow$ feeding sound only
2. Branches on the log $\Rightarrow$ feeding and delimbing sound combined

The sound of branches breaking was synthesized in the same fashion as the feeding sound itself. A real-life sound sample of a branch breaking was used as a basis for the synthesis process. First, the LPC filter modeling the spectral content was calculated from this sound sample using a filter order of $p = 1000$. This filter was then used to filter a white noise excitation, resulting in a noise-like signal with the correct spectral characteristics. Next, full-wave rectification and a sliding average filter with a window size of $N = 100$ were utilized to obtain the amplitude envelope of the original signal, cf. Fig. 32. Lastly, the resulting spectral envelope and amplitude envelope were multiplied to obtain a synthetic sound sample, cf. Fig. 33.

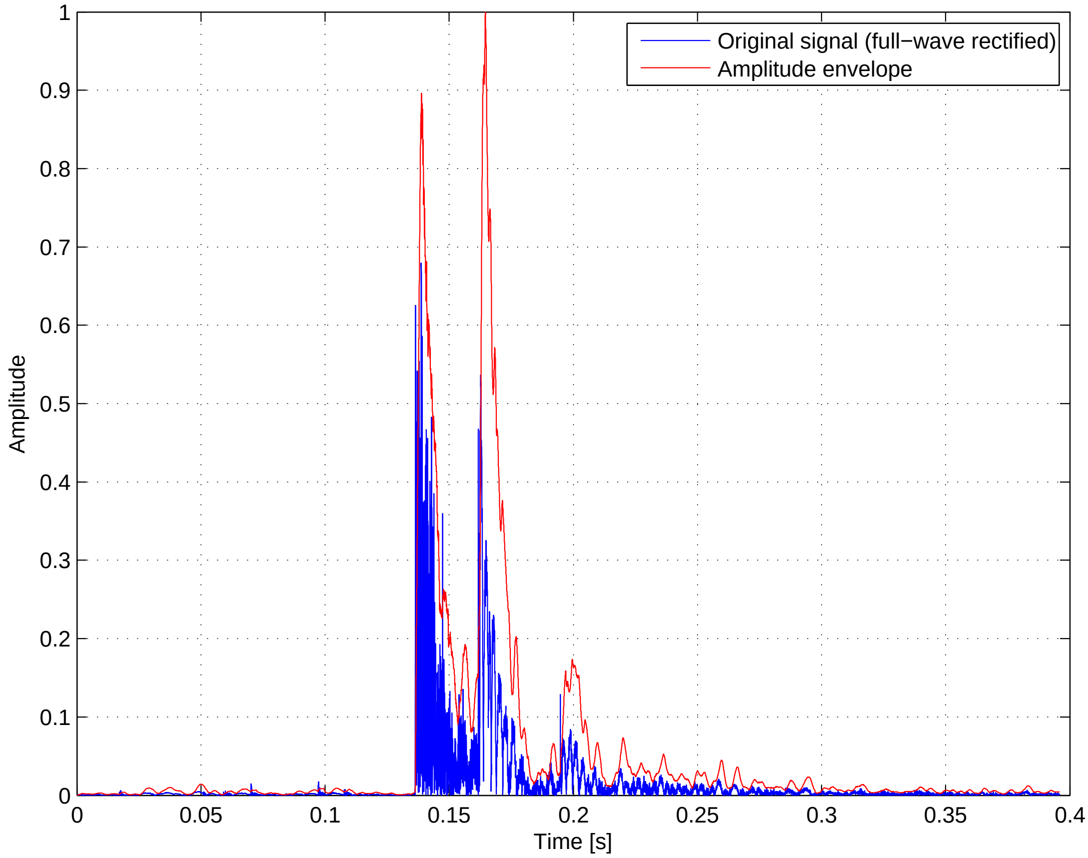

Figure 32: Time domain presentation of the original full-wave rectified signal of a branch breaking sound and its amplitude envelope (red).

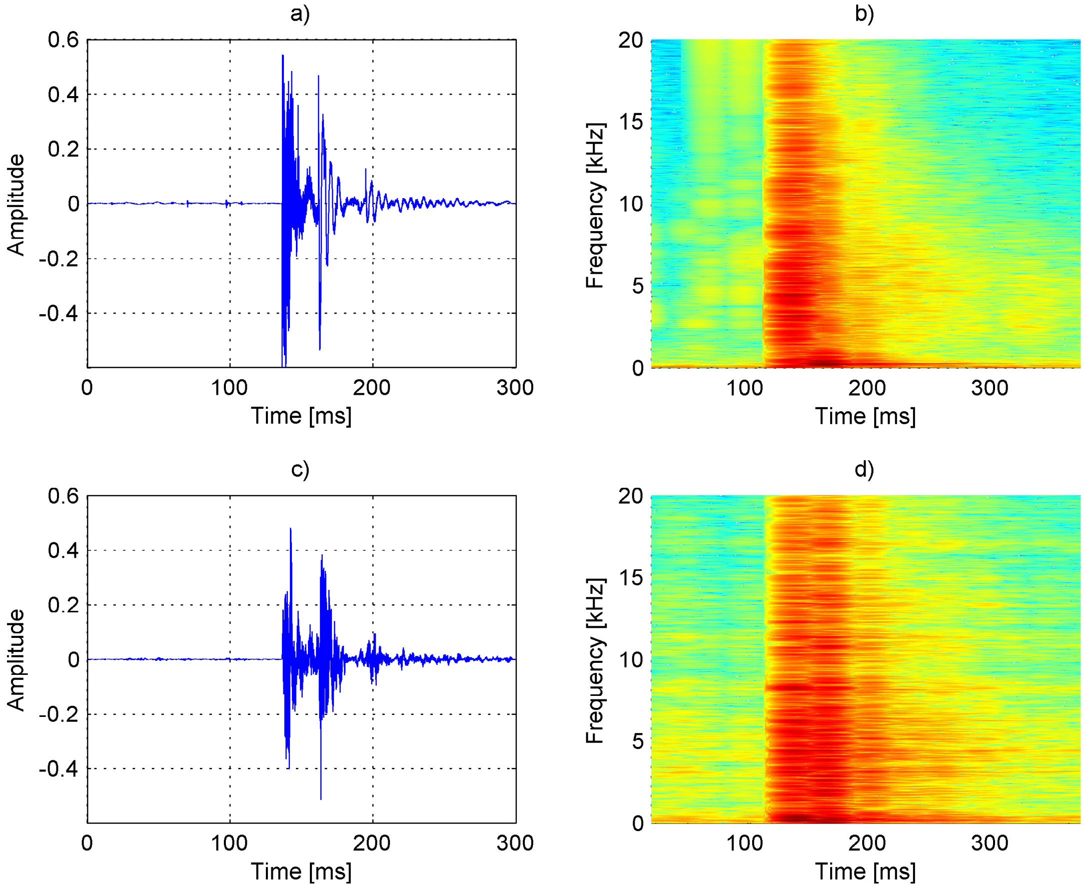

Figure 33: Time domain presentations and spectrograms of the a)b) original branch breaking sound and c)d) the synthetic signal created from it.

Several different real-life branch sounds were used to synthesize different versions of branch-breaking sounds adding variation and realism to the final combined sound of feeding and delimbing. Different synthetic branch-breaking sounds were chosen at random and spaced according to the peaks of the amplitude envelope of a real-life delimbing sound. The location of the peaks and their amplitudes were calculated using peak detection as explained in Sec. 2.5. A minimum peak height of 0.4 and a minimum peak separation of 1400 samples were used for the peak detection algorithm. The minimum height difference was set to zero. The detected peaks are shown in Fig. 34 below. Figure 35 presents a sample of several different delimbing sounds and the combination of the feeding and delimbing sound in the time domain.

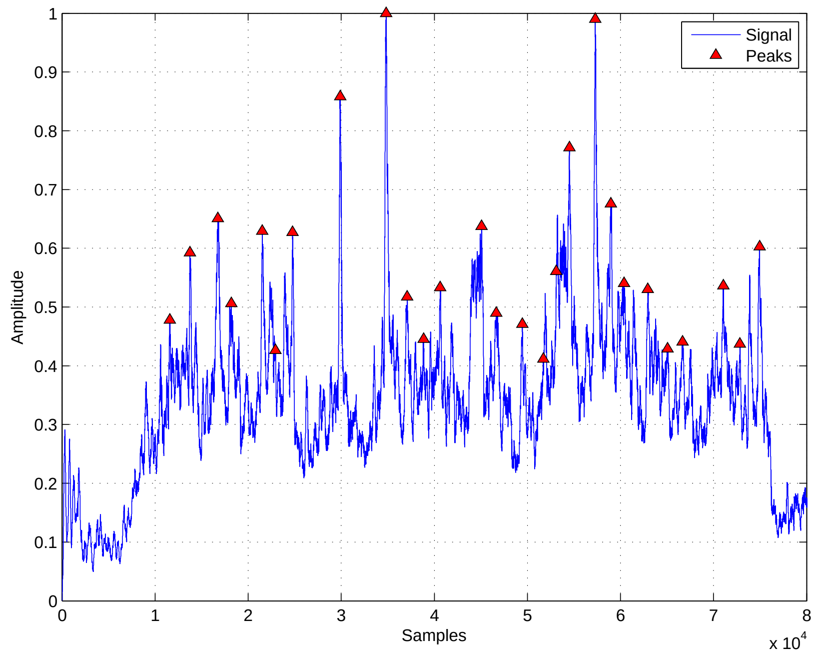

Figure 34: Peak detection of the amplitude envelope of a real delimbing sound.

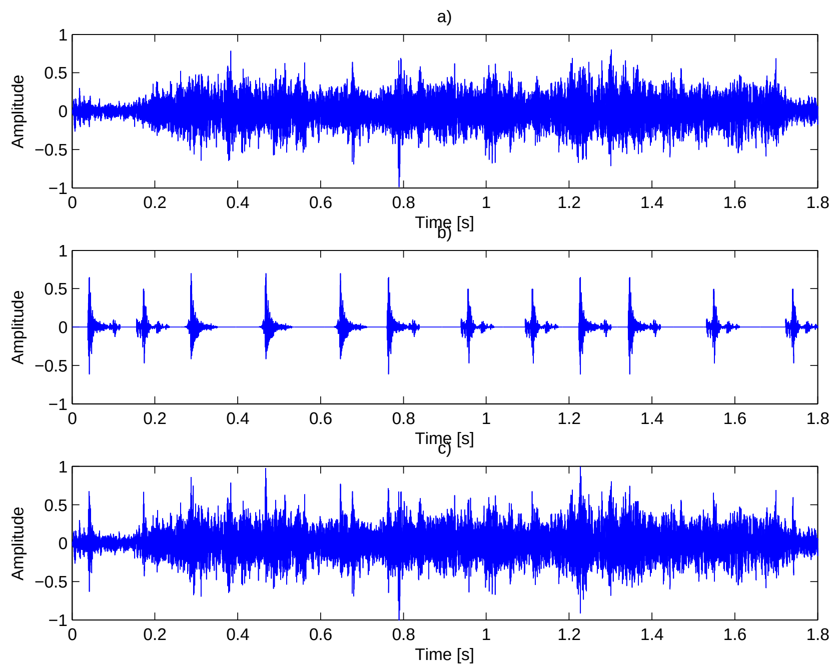

Figure 35: Time domain presentations of a) synthetic feeding sound only, b) delimbing samples, and c) feeding and delimbing combined.

5.4 Contact Sounds

The work related to the sounds of the forwarder simulator was mostly based on contact sounds, i.e. sounds caused by logs or parts of the machine making contact with the forwarder. Basic contact sounds consisted of: logs hitting the screen, logs hitting the bunks, and motion of the grapple opening and stopping. These aforementioned sound events are a fundamental part of the soundscape for a forest forwarder operator, and as such are required in the simulator. For example, the sound of logs hitting the screen is caused by the operator as the logs are deliberately struck against the screen to ensure they are as far back on the carriage as possible. Hearing this sound notifies the operator of the position of the logs.

The sound samples of the events described below were extracted from a video of a forest forwarder at work. The sound samples are quite short, but consist of heavy background noise caused by the diesel engine and strong gusts of wind. Spectral subtraction was used to remove the excess noise, resulting in clean sound samples, cf. Sec. 2.4. In other words, no actual synthesis was required for the contact sounds, as the processed sound samples were already of good enough quality.

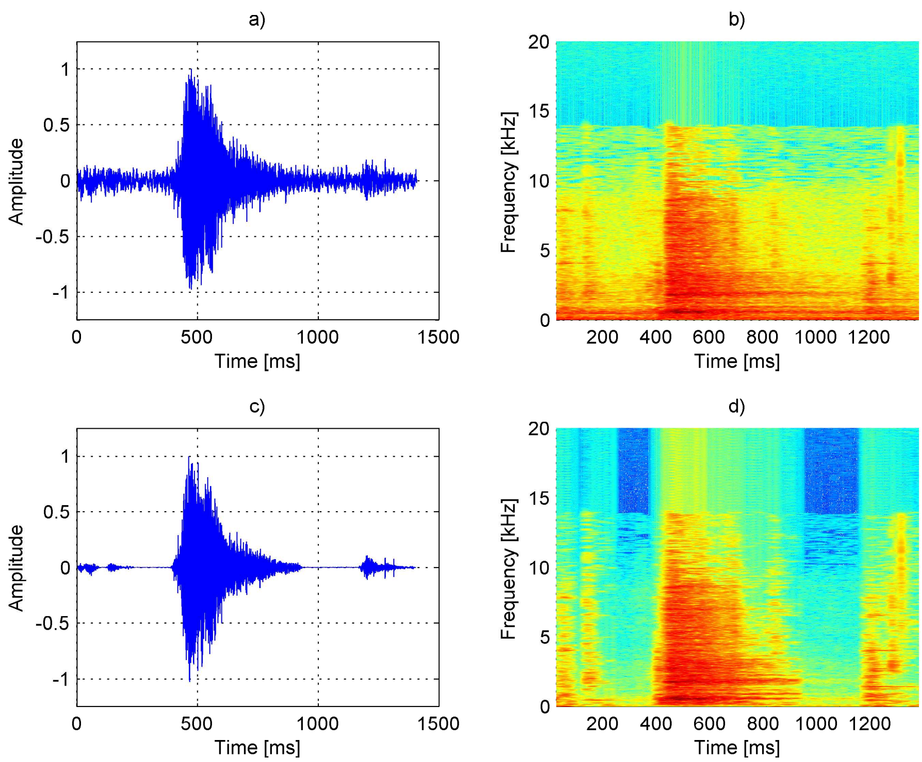

Figure 36: Time domain presentations and spectrograms of a)b) a noisy sample and c)d) a clean sample of the sound of logs hitting the screen.

5.4.1 Logs Hitting the Screen

As explained above, the sound of the logs hitting the screen gives vital information to the operator, and is thus a very important sound to have in the simulator. Figure 36 presents the original and processed sound samples of this event. As can be noticed from both the time domain and spectrogram presentations, there is significantly less noise present in the processed sound samples. This is especially evident in the spectrogram, where only the sound of interest is amplified and everything else is attenuated.

5.4.2 Logs Hitting the Bunks

The sound of the logs hitting the bunks is also a common event while using the forwarder. The bunks hold the logs on the carriage from both sides and contact with them is inevitable. The ringing-like sound of the bunks gives the operator important cues about the position of the grapple and logs, a valuable aid compared to visual observation only. Recognizing this sound already in the simulator phase of training can help future operators be well prepared for use of the actual machine. As with the screen contact sound, spectral subtraction succeeds in removing the background noise from the bunk contact sound sample, as is apparent in Fig. 37. Both the time domain and spectrograms of the processed signal show a significant decrease in noise.

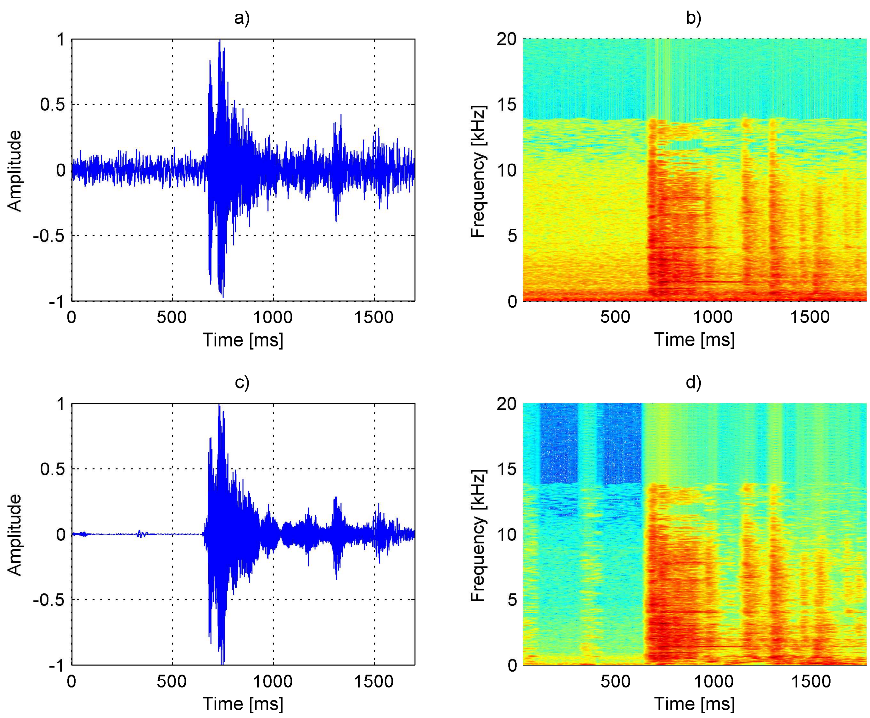

Figure 37: Time domain presentations and spectrograms of a)b) a noisy sample and c)d) a clean sample of the sound of logs hitting the bunks.

5.4.3 Grapple Opening

The grapples is the key component in the forwarder which allows it to operate as it does. It is as vital to the forwarder as the harvester head is to the harvester. The grapple is a claw-like tool which is used to grab and release logs, so they can be lifted off the ground and onto the forwarder (and vice versa). The opening of the grapple causes a sound event which should be very familiar to experienced operators. In addition to the opening itself, an additional sound is heard as the grapple reaches its maximum opened position. Although the opening of the grapple might not be considered a contact sound, there is still some contact in the grapple as it reaches its maximum. Again, spectral subtraction is successful in removing the background noise from the noisy sample. Figure 38 shows the noisy and noiseless signals in the time domain and their corresponding spectrograms.

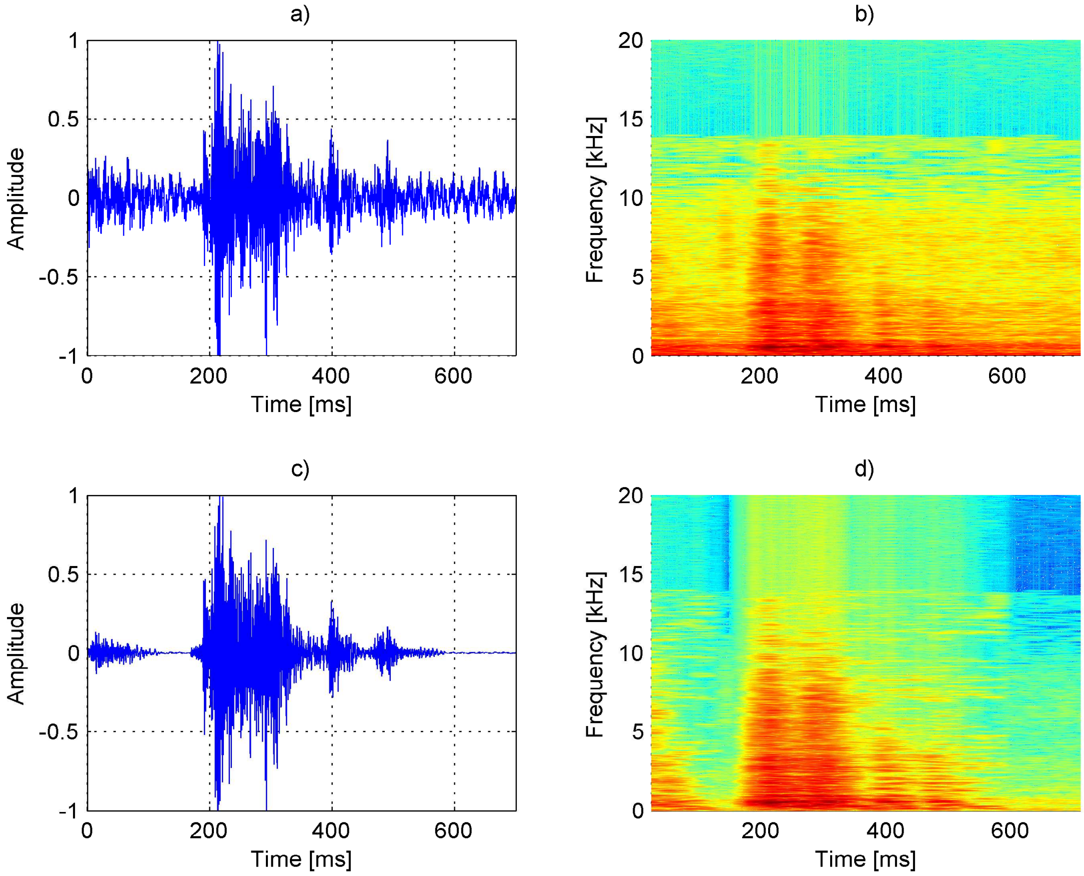

Figure 38: Time domain presentations and spectrograms of a)b) a noisy sample and c)d) a clean sample of the sound of logs the grapple opening and stopping.

5.5 Hydraulic Pump Sound

The hydraulic pump sound is a high frequency whistle-like sound found in the harvester. It is quite audible right before and after sawing, feeding, or delimbing and it is caused by the hydraulics of the machine. The sound can quite easily be mistaken for the sound of the turbo which is also a similar high frequency whistle, although at a higher frequency. The hydraulic pump sound can be heard in the range of 8.6 kHz to 10.5 kHz as seen in Fig. 39. The sound is visible in the spectrogram as a slim darker descending line in the aforementioned frequency range.

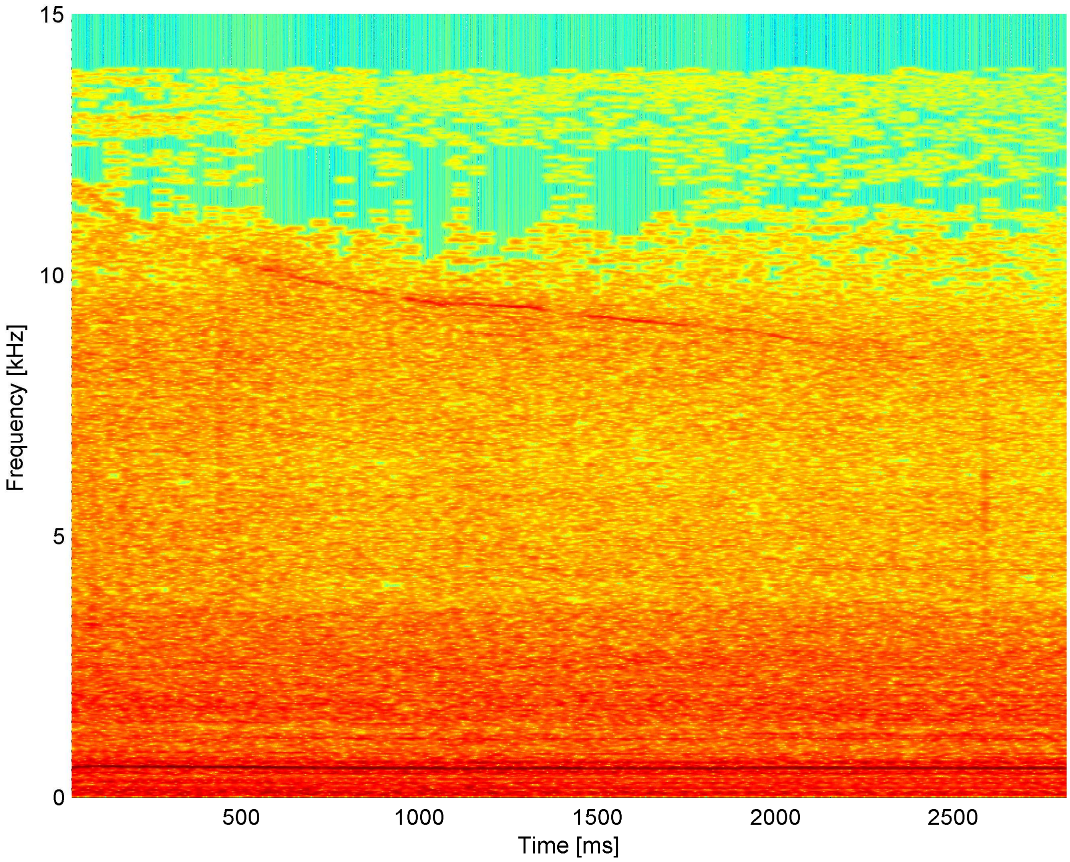

Figure 39: Spectrogram of the hydraulic pump sound. The sound itself is the dark red line around 8.6 kHz - 10.5 kHz.

The sound of the hydraulic pump was synthesized by first implementing a FIR band-pass filter at 8.6 - 10.5 kHz. This filtered output was used as the input for an LPC filter (order, p = 1000), which was then used to filter a white noise excitation, as was done with other hydraulic sounds explained in Sec. 3. To add some variation in frequency to the synthetic version, a variable state filter was utilized, cf. 2.3.4. The state variable filter alters the band-pass of the filter, causing the frequency to decrease and increase. Figure 40 presents the different steps of the synthetic process.

Although in the synthetic version the behavior of the darker line does not exactly replicate that of the original signal, it still sounds perceptually very similar.

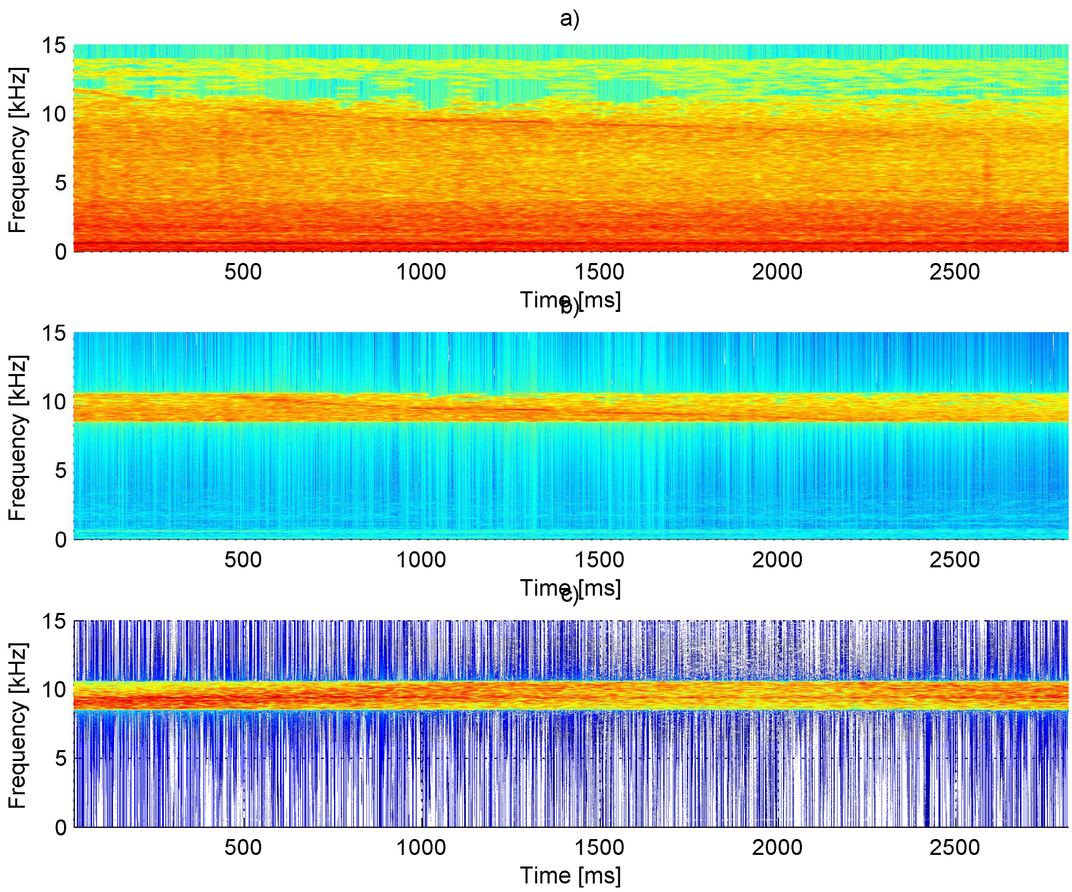

Figure 40: Spectrograms of the different synthesis steps. a) the original signal, b) the band-pass filtered signal, c) the synthetic signal filtered with a state variable filter.

5.6 Load Brake Sound

The load brake is a hinge-type component connecting the grapple to the boom of the forwarder. As its name implies, it operates as a brake preventing the grapple from swinging too violently as the boom is shifted around. The sound of the load brake can serve as an audible cue to the operator informing of too abrupt and reckless movements. This is why the sound of a load brake could serve as a valuable asset in a simulator, enabling the trainee to properly distinguish proper boom movements though audio cues via the load brake sound.

The sound of the load brake is practically inaudible in the forwarder videos used in previous sections of this chapter, thus a more vague approach had to be utilized for the synthesis. According to experts working with the simulator and machines, the sound is similar to a squeaky bicycle saddle, thus a sound clip of a bicycle saddle squeaking was used as the source material and basis for the synthesis.

The sound was synthesized in the same manner as the feeding sound. First, LPC (p = 1000) was used to calculate the spectral characteristics of the original sound which was then used to filter a white noise excitation. Next, the original signal was full-wave rectified in the time domain and it was filtered with a sliding average filter with a window size $N = 200$, cf. Fig. 41.

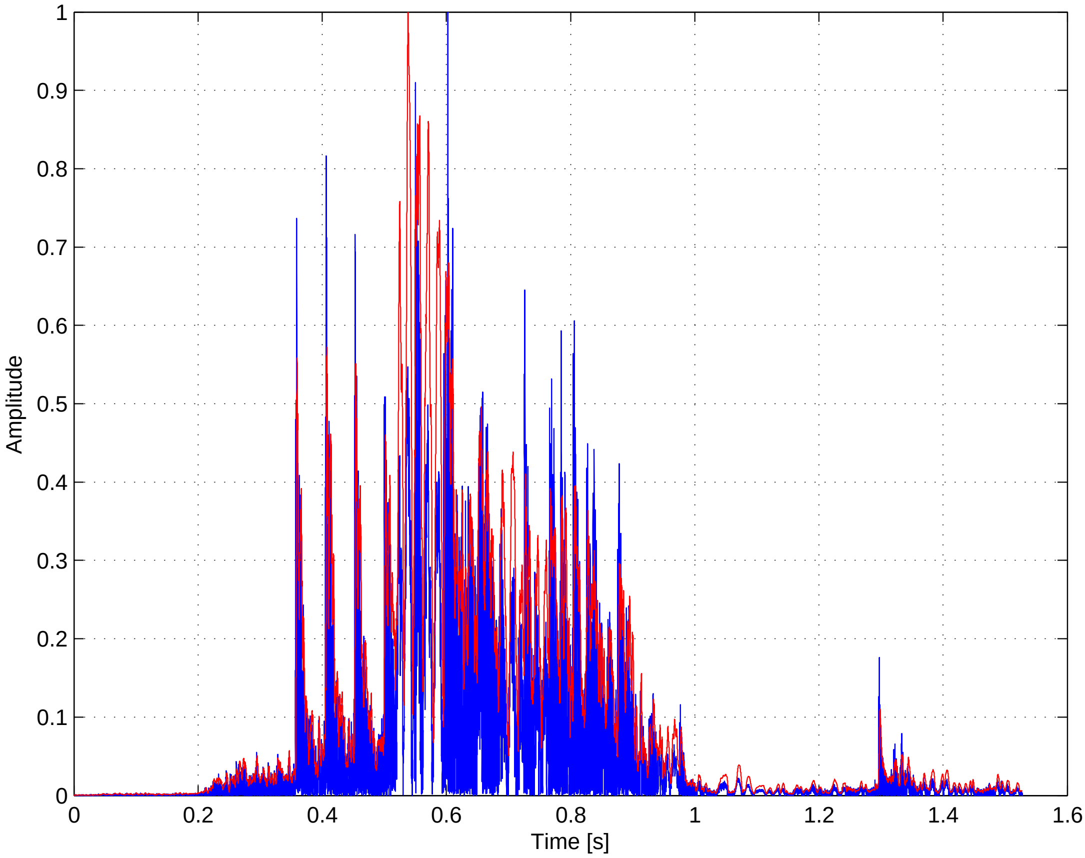

Figure 41: Time domain presentation of the original full-wave rectified signal and its amplitude envelope (red).

The LPC filtered excitation signal was then multiplied with the amplitude envelope to achieve the final synthetic result which is presented in Fig. 42. Even though the reference sound sample is not an actual load brake sound, the spectrograms and time domain presentations clearly show how the synthetic version resembles the reference sound. By applying this method to an actual load brake sound, the result would be of much higher quality.

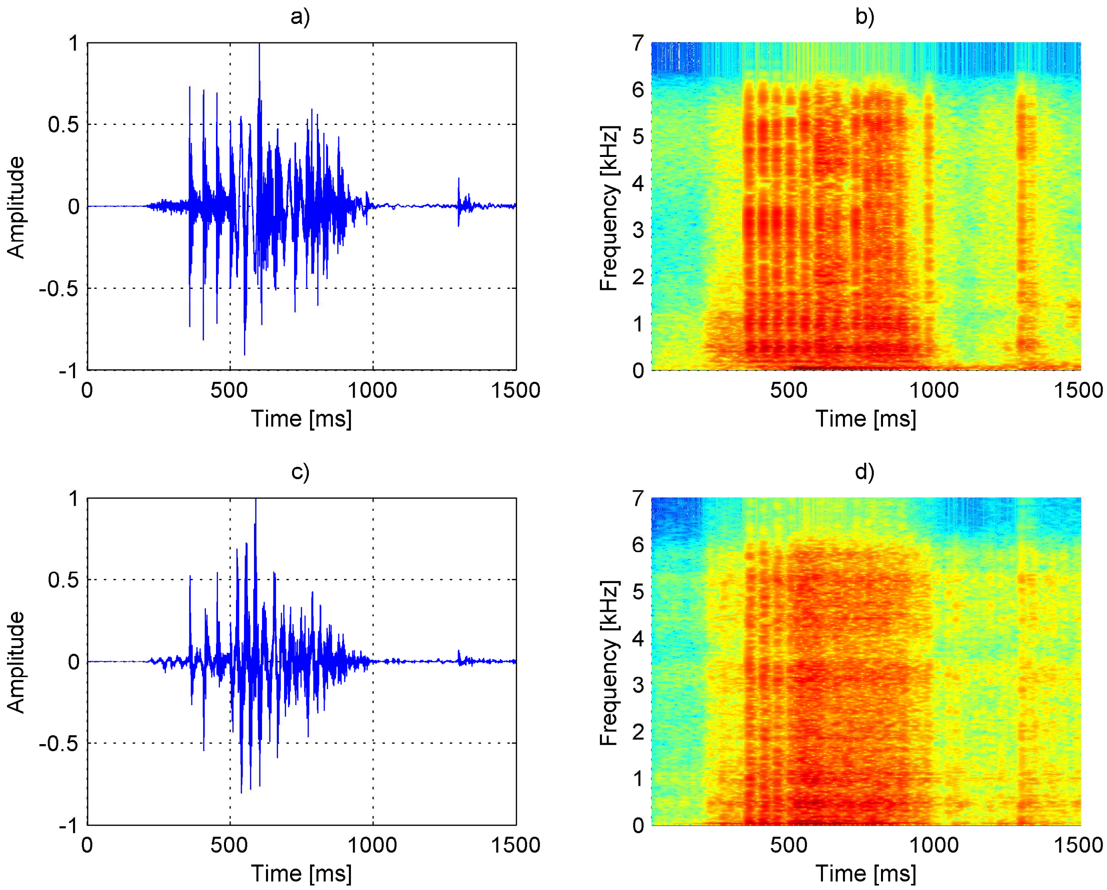

Figure 42: Time domain presentations and spectrograms of the a)b) original signal and c)d) synthetic signal.

6 Synthesis Evaluation

6.1 Listening Test

The hydraulic sounds and the harvester and forwarder sounds created in this project were evaluated in an informal listening test. Two experts were asked to first determine whether the sound sample sounds realistic or not, after which they could rate the sound on a scale of 0 to 5 (0 = poor, 5 = excellent). The expert listeners were both employees of the simulator company and both had excellent knowledge regarding the forwarder and harvester simulators. In addition, one expert had real life experience in operating the actual machines. Some of the sounds presented were played along with video of the actual event to more closely simulate how the sound would work in a simulator. The listening test itself was a very informal event where the experts could give comments and feedback orally and in writing at any time. The experts were also allowed to hear the samples as many times as required. The evaluation form can be seen in Appendix A and the answers given by the experts in Appendix B.

The idea behind the listening test was to receive feedback from people with real life experience with the machines and simulators in question. Their feedback is essential in determining whether the sounds created for the simulators actually correspond to the actual machines they are simulating. The feedback and results were fairly positive, although some sounds were deemed unrealistic and requiring improvements. Though some sounds failed in sounding realistic enough, the listening test as a whole was very successful as it lead to some excellent discussion and improvement ideas for the sounds. The results and improvements will be discussed below.

6.2 Results and Improvements

Table 1 shows the results of the listening test. The test Experts are labeled as A and B, where A is the test Expert with experience in operating the actual machines.

For the first question regarding whether the sound samples actually sound like the sounds they are simulating, Expert A answered "Yes" 8 out of 9 times (89 %) and Expert B 6 out 9 times (67 %). The load brake sound was the one sound both experts agreed did not sound realistic. Both experts still considered a majority of the sounds to correlate to their corresponding sound events. The calculated average scores regarding the authenticity of the sounds were very similar for both experts. Expert A rated the sounds with an average of 3.1 and Expert B with an average rating of 3.0.

6.2.1 Feeding

The feeding sound was played along with a video clip of feeding and it received fairly positive feedback from both experts. Expert A suggested amplifying the sound of the feed rollers, which can apparently be a very loud sound due to the rollers scraping against the bark of a tree. Expert B noted that the authentic characteristics of

Table 1: The results of the listening test.

	Does the sound sample sound like the sound event in question?		How realistic is the sound?	
Sound Event	**Expert A**	**Expert B**	**Expert A**	**Expert B**
Feeding	Yes	Yes	3	3
Feeding & Delimbing	Yes	No	2	0
Hydraulic Sounds	Yes	Yes	4	5
Hydraulic Pump	Yes	Yes	4	4
Contact Bunks	Yes	Yes	4	5
Contact Screen	Yes	No	3	2
Contact Grapple	Yes	Yes	4	4
Contact Logs	Yes	Yes	4	3
Load Brake	No	No	0	1

the sound had been found, although the result could have been much better with higher quality source material. Expert B's point is a very valid one, as the source material was littered with engine and wind noise, which however was suppressed by filtering, will surely cause loss of quality in the final synthetic result. With a cleaner recording of a feeding event, the loud sound of the feed rollers could possibly be more prominent in the final result.

6.2.2 Feeding & Delimbing

The delimbing sound was played combined with the feeding sound (along with a video clip), but it did not perform well in the listening test. The single crackling sounds were also played separately to the experts. Expert A gave the sound a rating of 2.0 and Expert B did not consider it realistic at all and rated it a 0.0. According to the discussion and feedback on the delimbing sound, the synthesis method itself is fine, but the "crack" or "snap" sounds created are too short and clean sounding. This is mostly due to the source material used, meaning more appropriate sounds of branches breaking would need to be either recorded or obtained. Expert A commented that the sounds more closely resembled the sound of an ax chopping wood, instead of the delimbing knives cutting branches.

6.2.3 Hydraulic Sounds

The hydraulic sounds were played back in a longer sound clip with a combination of the different hydraulic sound types presented in Sec. 3. This longer sound clip aimed to simulate the sound of hydraulic cylinders at work. The single sounds were also played separately and compared to the real hydraulic sounds. The feedback for the sounds was excellent, as Expert A gave the sounds a 4.0 rating and Expert A a rating of 5.0. Expert A especially praised the high frequency variable hydraulic sound, which sounded extremely realistic, cf. Sec. 3.3.4.

6.2.4 Hydraulic Pump

The hydraulic pump sound of the harvester also received positive feedback, which is not surprising as it was synthesized in the same fashion as the other hydraulic sounds, which also received excellent reviews. The sound was played along with a video clip similar to the other harvester and forwarder sounds. Most of the discussion on this particular sound was in determining the difference between the squeal of the turbo and the whistling sound created by the hydraulic pump. According to both experts, both sounds are very similar as they are both high frequency "squeals" or "whistling sounds". In the end, both experts agreed the sound of the turbo is an even higher frequency sound and this particular synthetic sound is that of the hydraulic pump.

6.2.5 Forwarder Contact Sounds

Four different contact sounds were played back individually to the experts and each sound was rated separately. The sounds were also played back with a short video clip where they were all utilized to present the sounds in the correct context. The four contact sounds reviewed were: logs hitting the bunks (vertical poles), logs hitting the screen (the back of the carriage), sound of the grapple opening, and the sound of logs dropping. As shown in Table 1, the contact sounds scored fairly well with both experts. The screen sound received some critique as the sound of the grapple opening can apparently be heard in the sound sample. A lower frequency sound was also requested for the sound of the logs dropping, as they currently sounded a bit too synthetic.

6.2.6 Load Brake Sound

The synthetic load brake sound scored very poorly in the test (A: 0, B: 1). This is mostly due to a lack of source material, as the load brake is extremely hard to hear in the forwarder videos (practically inaudible). The sound was synthesized on the basis of the sound sounding like a squeaky bicycle saddle. The synthesis method itself is correct, but an actual recording of a load brake would be required to achieve a realistic result.

7 Conclusions and Future Work

This thesis presented different sound synthesis methods for working machine simulators. Sounds were synthesized for three different simulators and their corresponding machines: the forest harvester and forwarder, a drill rig, and a truck-mounted hydraulic platform. Several different signal processing techniques and synthesis approaches were employed, including filtering, spectral subtraction, linear predictive coding, and peak detection. The thesis was divided into four main parts: hydraulic sounds, drilling sounds, forwarder/harvester sounds, and evaluation.

Hydraulic sounds were synthesized using LPC and a white noise excitation signal. Several different hydraulic sounds were first analyzed and extracted from the source material and then synthesized. These hydraulic sounds included: the basic sound, a fading sound, a high frequency sound, a variable high frequency sound, and three different piston contact sounds. The benefits of the synthetic versions of these sounds include the ability to create sound samples of any length required without a looping sound effect and the advantage of being completely free of background noise. The hydraulic sounds created in this project can be utilized in any of the three simulators, as the machines themselves rely heavily on hydraulic operations. The simulators did not include existing hydraulic sounds, thus these newly created sounds can add a new element of reality to the training process.

A simple drilling sound already existed in the drill rig simulator which utilized pitch shifting to alter the drilling frequency. This thesis aimed to improve the drilling sound by removing the need for pitch shifting, which can often lead to very unrealistic sounding results. This was achieved by separating single click sounds from the original drilling sound sample and playing them back at the correct drilling frequency. Statistical analysis was applied to synthesize drilling sounds correlating to different drilling scenarios: normal drilling, underfeed, and overfeed. Using a sliding average filter and peak detection, statistical data was obtained from real drilling sounds regarding the separation between the clicks and their amplitude variations. Although the drilling sound synthesis method is an improvement to the existing approach, improvements are still required to achieve a truly realistic drilling sound.

The main sounds synthesized for the forest harvester and forwarder were feeding, delimbing, and basic contact sounds. LPC and sliding average filters were employed in the synthesis of the feeding and delimbing sounds. LPC was used to extract the spectral features of the sounds and sliding average filters isolated the amplitude envelopes, which combined with a white noise excitation formed a synthetic result. Contact sound samples were simply created by applying spectral subtraction to noisy real life sound samples of the different contact situations.

The last section presented a listening test which was performed regarding the hydraulic and forest machine sounds. Two test Experts were employed in evaluating the sounds created in a very informal listening test, which consisted of two simple questions and oral feedback. The sounds scored fairly well in the evaluation with some sounds performing better than others. The session was an extremely fruitful one with excellent discussion and ideas on how to improve the sounds.

Future work on the Expert should prioritize on the drilling sound, as it is still quite synthetic sounding. The most promising solution would be to obtain a clean sample of a real life click sound and utilize it in the same fashion as the synthetic clicks were in this project. By combining statistical analysis with different clean samples of different drilling sound types, all the variations of drilling could be synthesized in a realistic way. The rattling sounds would especially benefit from this, as the sound characteristics of the synthetic drilling sound are immensely different from an actual rattling sound.

Better source material would also help in improving some of the harvester and forwarder sounds. Delimbing would require better and more suitable recordings of branches breaking to correctly mimic the sound of a delimbing event. The sound of a load brake could also be successfully synthesized with an actual recording of the sound. Although in some cases the lack of proper source material decreased the quality of the synthetic results, the methods presented in this thesis should still prove to be valuable and useful in improving the sound environment in working machine simulators.

References

[1] A. Okapuu-von Veh, R. J. Marceau, A. MAlowany, P. Desbiens, A. Daigle, R. Gauthier, A. Shaikh, and J. C. Rizzi, "Design and Operation of a Virtual Reality Operator-Training System," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1585-1591, Aug. 1996

[2] H. Ploner-Bernard, A. Sontacchi, G. Lichtenegger, and S. Vössner, "Sound-System Design for a Professional Full-Flight Simulator," *Proc. of the 8th Int. Conference on Digital Audio Effects (DAFx'05)*, Madrid, Spain, Sept. 20-22, 2005

[3] S. Vössner, R. Braunstingl, H. Ploner-Bernard, and A. Sontacchi, "A New Functional Framework for a Sound System for Realtime Flight Simulation," *Proc. of the 8th Int. Conference on Digital Audio Effects (DAFx'05)*, Madrid, Spain, Sept. 20-22, 2005

[4] S. Oksanen, J. Parker, and V. Välimäki, "Physically Informed Synthesis of Jackhammer Tool Impact Sounds," *Proc. of the 16th int. Conference on Digital Audio Effects (DAFx-13)*, Maynooth, Ireland, Sept. 2-5, 2013

[5] S. Oksanen, J. Parker, and V. Välimäki, "Vibroacoustic Analysis and Synthesis of Struck Metal Bars Using Musical Instrument Modeling Techniques," *Akustiikkapäivät 2013*, Turku, Finland, May 22-23, 2013

[6] P. R. Cook, *Real Sound Synthesis for Interactive Applications*, AK Peters, Ltd., 1st edition, 2002.

[7] D. O'Shaughnessy, "Linear Predictive Coding," *IEEE Potentials*, vol. 7, pp. 29-32, Feb. 1988.

[8] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, pp. 561-580, Apr. 1975.

[9] J. O. Smith, *Viewpoints on the History of Digital Synthesis*, `https://ccrma.stanford.edu/~jos/kna/`, accessed June 26, 2014, online book.

[10] J. F. Kaiser, "On the Use of the I_0-Sinh Window for Spectrum Analyis," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no. 1, pp. 105-107, Feb. 1980.

[11] E. W. Weisstein, "Modified Bessel Function of the First Kind," From *MathWorld–A Wolfram Web Resource*, `http://mathworld.wolfram.com/ModifiedBesselFunctionoftheFirstKind.html`, accessed October 16, 2014, online book.

[12] J. O. Smith, *Introduction to Digital Filters*, `https://ccrma.stanford.edu/~jos/filters/`, accessed July 14, 2014, online book.

[13] S. Golestan, M. Ramezani, J. M. Guerrero, F. D. Freijedo, and M. Monfared, "Moving Average Filter Based Phase-Locked Loops: Performance Analysis and Design Guidelines," *IEEE Transactions on Power Electronics*, vol. 29, no. 6, pp. 2750-2763, Jun. 2014

[14] R. A. Moog, "A Voltage-Controlled Low-Pass High-Pass Filter for Audio Signal Processing," *17th AES Convention*, New York, Oct. 1965.

[15] A. Huovilainen, "Non-Linear Digital Implementation of the Moog Ladder Filter," in *Proc. 7th Intl. Conf. Digital Audio Effects*, Naples, Italy, Oct. 2004, pp. 61-64

[16] V. Välimäki, A. Huovilainen, "Oscillator and Filter Algorithms for Virtual Analog Synthesis," *Computer Music Journal*, vol. 30, no 2. Summer 2006, pp. 19-31

[17] R. G. Lyons, *Understanding Digital Signal Processing*, Pearson Education, Inc., 3rd edition, 2011, pp. 178-196

[18] U. Zölzer, *DAFX - Digital Audio Effects*, John Wiley & Sons, Ltd., 2002, pp. 35-36

[19] S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-27, no 2., pp. 113-120, Apr. 1979.

[20] J. B. Allen, "Short Time Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-25 (3): 235-238, Jun. 1977.

[21] H. F. G. Nia, H. Hu, "Applying Bayesian Decision Theory to Peak Detection of Stochastic Signals," *2012 4th Computer Science and Electronic Engineering Conference (CEEC)*, pp. 117-122, Sep. 2012

[22] M. Heiniö, *Rock Excavation Handbook* Sandvik Tamrock Corp. 1999

[23] NIST/SEMATECH e-Handbook of Statistical Methods, `http://www.itl.nist.gov/div898/handbook/`, accessed September 26, 2014, online book.

A Appendix

The evaluation form used in evaluating harvester and forwarder sounds.

REMES
Harvester/Forwarder Sound Evaluation

1. Feeding

a) Does the sound sample sound like feeding? Yes [ ] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

0 1 2 3 4 5

c) Comments/Improvements:

2. Feeding & Delimbing

a) Does the sound sample sound like feeding and delimbing? Yes [ ] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

0 1 2 3 4 5

c) Comments/Improvements:

3. Hydraulic cylinders

a) Does the sound sample sound like actual hydraulic sounds? Yes [ ] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

0 1 2 3 4 5

c) Comments/Improvements:

4. Turbo/hydraulic pump whistling

a) Does the sound sample sound like the whistling of the turbo or hydraulic pumps?

Yes [ ] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

0 1 2 3 4 5

c) Comments/Improvements:

5. Forwarder contact sounds

a) Do the contact sound samples sound like actual contact sounds?

Logs hitting the bunks (the vertical poles)	Yes [ ]	No [ ]
Logs hitting the screen (the back of the carriage)	Yes [ ]	No [ ]
Sound of the arm opening	Yes [ ]	No [ ]
Sound of logs dropping	Yes [ ]	No [ ]

b) How realistic do the samples sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

Logs hitting the bunks (the vertical poles)

0 1 2 3 4 5

Logs hitting the screen (the back of the carriage)

0 1 2 3 4 5

Sound of the arm opening

0 1 2 3 4 5

Sound of logs dropping

0 1 2 3 4 5

c) Comments/Improvements:

6. Load brake sound (riipukejarru)

a) Does the sound sample sound like the actual load brake?

Yes [ ] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

0 1 2 3 4 5

c) Comments/Improvements:

7. Other comments/suggestions/improvements

B Appendix

The evaluation form with the answers of the experts typed out. Expert A's answers were translated from Finnish to English.

REMES
Harvester/Forwarder Sound Evaluation

1. Feeding

a) Does the sound sample sound like feeding? Yes [**AB**] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

A: 3

B: 3

c) Comments/Improvements:

A: "More sound from the feeding rollers, a strong sound from the bark of the tree"

B: "Based on the source material that wasn't of the best quality, the authentic characteristics of the sound were found. Avg. score because of the source."

2. Feeding & Delimbing

a) Does the sound sample sound like feeding and delimbing? Yes [**A**] No [**B**]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

A: 2

B: 0

c) Comments/Improvements:

A:

B: "Too artificial cracks (too short, too 'clean')"

3. Hydraulic cylinders

a) Does the sound sample sound like actual hydraulic sounds? Yes [**AB**]No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

A: 4

B: 5

c) Comments/Improvements:

A: “High flow sound is good.”

B: “Sounds great.”

4. Turbo/hydraulic pump whistling

a) Does the sound sample sound like the whistling of the turbo or hydraulic pumps?

Yes [**AB**] No [ ]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

A: 4

B: 4

c) Comments/Improvements:

A: “Hydraulic pump sound, turbo is higher”

B: “It’s not the turbo sound, but the pump sound. Quite good at that.”

5. Forwarder contact sounds

a) Do the contact sound samples sound like actual contact sounds?

Logs hitting the bunks (the vertical poles)	Yes [**AB**]	No [ ]
Logs hitting the screen (the back of the carriage)	Yes [ ] **A**	No [**B**]
Sound of the arm opening	Yes [**AB**]	No [ ]
Sound of logs dropping	Yes [**AB**]	No [ ]

b) How realistic do the samples sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

Logs hitting the bunks (the vertical poles)

A: 4 B: 5

Logs hitting the screen (the back of the carriage)

A: 3 B: 2

Sound of the arm opening

A: 4 B: 4

Sound of logs dropping

A: 4 B: 3

c) Comments/Improvements:

A:

B: "Screen sound contains grapple opening 'cling'. Log drop sounds need a lower frequency."

6. Load brake sound (riipukejarru)

a) Does the sound sample sound like the actual load brake?

Yes [ ] No [**AB**]

b) How realistic does the synthetic sample sound? Circle the appropriate value.
(0 = poor, 5 = excellent)

A:

B: 1

c) Comments/Improvements:

A:

B: "Too distorted (maybe due to speaker setup)."

7. Other comments/suggestions/improvements

A:

B: "Nice analysis/synthesis, some sounds were excellent and some unrecognizable. However the results are clearly very useful for training purposes."